Marriage, Fertility, and Cultural Integration in Italy*

Alberto Bisin† Giulia Tura‡

April 6, 2024

Abstract

We study cultural integration as an equilibrium of marital matching along cultural and education lines and intrahousehold decisions regarding fertility and cultural socialization. Structural estimates reveal strong demand to preserve cultural identity on the part of immigrants and little acceptance of the immigrants' cultural diversity of natives. Also, these estimates reveral that less-educated parents care relatively (slightly) more about the cultural identity of their children and that the parental value of children's education is heterogeneous across cultures. Nonetheless, we simulate substantial, though heterogeneous, integration rates across immigrant groups, 75% on average over one generation. Counterfactuals show how more accepting preferences of the natives would lead to slower cultural integration, while a reduction in economic incentives to immigrants would increase it. Finally, we evaluate a policy enhancing social welfare by strengthening the ethnic network of immigrants.

JEL Codes: D1, J12, J13, J15.
Keywords: Marital Matching, Fertility, Cultural Transmission, Integration.

*We thank Jérôme Adda, Pierre-André Chiappori, Alfred Galichon, Yujung Hwang, Ilse Lindenlaub, Corinne Low, Costas Meghir, Paolo Pinotti, Jean-Philippe Platteau, Marco Tabellini, Simon Weber and participants in seminars at many universities and conferences for very useful comments and discussions. Special thanks to Fulvio Ortu for support. Data used in this paper has been examined at the ADELE Laboratory (*Laboratorio per l'Analisi dei Dati ELEmentari*) at ISTAT, in compliance with the laws on the protection of statistical confidentiality and of personal data. We are solely responsible for the results and the opinions expressed in this paper, which do not constitute official statistics. We are grateful to Lorena Viviano and Silvia da Valle of the ADELE Laboratory in Milan and Florence for their support in providing us useful documentations and suggestions.

†Department of Economics, New York University, 10012 New York, USA. email: alberto.bisin@nyu.edu.
‡Department of Economics, Management and Statistics, University of Milano-Bicocca, 20126 Milan, Italy. email: giulia.tura@unimib.it.

1 Introduction

The recent surge in migration flows into Western countries represents one of the most contentious political and socio-economic phenomenon of the last decades. Widespread restrictive immigration policies are motivated in part by the perceived cultural externalities immigration imposes on natives in the integration process.[1] Indeed, the empirical evidence documents slow cultural integration on the part of minorities.[2]

To better understand the political economy of immigration, we study the dynamics of cultural integration, as the outcome of a process of marriage formation, fertility, and intra-household cultural decisions. In this context, cultural integration is an equilibrium phenomenon. On the demand side, immigrants trade off economic incentives to integrate, e.g., in the labor market, with preferences for preserving their cultural identity. On the supply side, natives modulate various degrees of (lack of) acceptance of immigrants' cultural-ethnic traits. Separating these demand and supply components at equilibrium is paramount to provide an adequate empirical basis for evaluating the dynamics of integration of immigrant minorities in a long term perspective and for assessing possible counterfactual interventions.[3]

To this end, we are able to exploit rich administrative marriage data providing information on the cultural-ethnic identity of immigrants in Italy. These micro-level data, provided by the Italian Statistical Institute (ISTAT, ADELE Laboratory), cover the universe of marriages formed in Italy from 1995 to 2012 and the universe of births and separations registered in the same period.[4] In addition, we recover a measure of parental cultural socialization of children from the *Condition and Social Integration of Foreign Nationals* Survey (2011-2012). Specifically, we interpret "speaking Italian at home" as a proxy for (negative) socialization rates to the culture of origin, that is, an indication that parents are not exercising much effort in transmitting their culture to their children.[5]

We document the relevance of cultural-ethic characteristics, as central determinants of the observed marriage, fertility, and socialization patterns. Indeed, we observe strong positive assortative mating along cultural-ethnic lines and sizable differences between homogamous and heterogamous marriages in terms of fertility and socialization, with homogamous mar-

[1] Negative labor market effects of immigration on natives are far from well-documented; see Borjas (2003), Card (1990), Dustmann et al. (2017), and Bisin and Zanella (2017).

[2] See e.g., Algan et al. (2012) and Abramitzky and Boustan (2017).

[3] See (Vieira et al., 2022) and (Fouka, 2024) for comments along these lines.

[4] Immigration to Italy has steadily increased over the past decades, with immigrants representing 10% of the total resident population in 2018.

[5] By exploiting further administrative data, in Section 5.3.2 we show that, in fact, immigrant students' who speak Italian at home display more social integration, e.g., in terms of ethnic identity and of achievement and educational choices.

riages displaying systematically higher childbirth investments. Finally and most importantly we take into account also educational attributes as determinants of marriage patterns - distinguishing high- (higher than high-school) from low-educated individuals. In this respect we document the presence of positive sorting by education, but also a much higher strength of sorting by culture than by education, even accounting for systematic differences in spouses' culture. We also show that the effects of parental education on fertility and socialization is systematically mediated by the cultural characteristic of the marriage.

This reduced form analysis, however, does not allow us to well identify the determinants of the observed process of cultural integration of immigrants. First of all, this analysis cannot allow for a decomposition of how much marriage gains are determined by the immigrants' preferences to preserve their cultural identity or by different forms of the natives' (lack of) acceptance of immigrants' cultural-ethnic traits - what we referred in the Introduction as the demand and the supply of integration. Furthermore, this analysis is silent about the underlying possible interactions between preferences along cultural and educational lines. It cannot separate preferences ranking spouses by education level from preferences valuing educational homogamy *per se* in marriage; and it does not uncover any systematic relationship between parents' culture and education and their ordering of the characteristics contributing to children's quality as well as their representation of the quantity-quality trade-off.

Our empirical study of cultural integration, therefore, relies on a structural model of the marriage market in which spouses match along the cultural-ethnic dimension as well as along the educational dimension and solve a collective household decision problem determining fertility and children socialization. Spouses match along cultural-ethnic and educational lines, identified by individuals' culture of origin and whether they have higher than high-school education. Marital utility depends on observable spouses' characteristics and results from a collective household decision problem embedded within this matching framework. Within marriage, parents choose fertility, investments in the cultural socialization of children, and possibly divorce.

More specifically, parents care about their children's education and the premium associated to a skilled child depends on the cultural-ethnic group of the parent. Parents also care about their children's cultural identity - and these preferences depend on the parents' original cultural-ethnic group and on their own education. With regards to transmission of education we simply postulate that highly educated parents have highly educated children with higher probability - which depends on the cultural-ethnic group and is calibrated from the data. Cultural socialization is instead the outcome of parental choice. Parents who value their cultural identity care about socializing their children, and they are endowed with socialization technologies to transmit their own cultural-ethnic traits to children. Thus, par-

ents choose to exert a direct socialization effort to affect their children's process of cultural identity formation. But effort requires parental resources, e.g., time spent with children, private school tuition, selection of residential neighborhoods and ethnic network and so on.

Socialization incentives vary then with the education and the cultural preferences of the parents. Furthermore, socialization technologies vary between culturally homogamous and heterogamous marriages. For instance, families where parents share the same cultural traits enjoy a more efficient socialization technology in transmitting their shared trait than families where parents do not share the same culture. Similarly, fertility and divorce choices depend on the education and the cultural preferences of the parents, as well as on the family type (homogamous or heterogamous along both the education and the cultural lines). In conclusion, the model implies a systematic dependence of marriage, fertility, socialization, and divorce patterns, across households' education and cultural-ethnic characteristics.

We estimate the parameters of the structural model exploiting variation across the cultural-ethnic composition of marital matches and across marriage markets. The main parameters of interest in the model are *cultural intolerance parameters* - a measure of the preference for the inter-generational transmission of culture of a specific cultural-ethnic group. The cultural intolerance of immigrants identifies the demand side of integration at equilibrium; specifically, whether economic incentives for integration dominate the preferences for cultural identity in driving the immigrants' marital, fertility, and intergenerational socialization choices. The cultural intolerance of natives, on the other hand, identifies the supply of cultural acceptance of the immigrants' cultural diversity in the society. Of great interest are also the preference parameters capturing the interactions between parental preferences along the cultural and the education line.

Estimated cultural intolerance parameters are positive, asymmetric, and highly heterogeneous across cultural-ethnic groups. In other words, preferences for cultural identity dominate economic (and legal) incentives to integration for all minorities. This is particularly so for immigrants from North Africa-Middle East, whose estimated cultural intolerance is more than four times as high as Europeans'. On the other hand, we also estimate high cultural intolerances for the Italian majority; that is, little cultural acceptance overall. In particular, Italians are the least accepting towards immigrants from Sub-Saharan Africa and North Africa-Middle East (estimates are twice as high as those towards immigrants from Europe). As for education, our estimates imply that parental cultural intolerances do not vary much with their education, but tend to be higher for unskilled immigrant parents.

We investigate the evolution of the distribution of the population by cultural traits in the long-run, by simulating our model at the estimated parameters over successive generations. Despite high cultural intolerance estimates, all cultural-ethnic minorities are simulated to

integrate to the Italian majority. The integration rate, defined as the reduction in the fraction of the total population (immigrants and natives) which is composed of immigrants who are not integrated to the native Italian culture, is 75% in one single generation. However, the pace of convergence is heterogeneous across cultural-ethnic groups. On the one hand, we find that minorities from Europe and North Africa-Middle East integrate rapidly. A slower integration rate characterizes instead immigrants from East Asia, Sub-Saharan Africa minorities, and especially Latin America, which exhibits an integration rate of 54% and reaches full convergence only after four generations. The patterns of cultural integration across ethnic groups in the simulations are not only the result of demand and supply effects at equilibrium; that is, of the cultural intolerance of immigrants towards the natives and of that of the natives towards the immigrants. Immigrants from North Africa-Middle East, Sub-Saharan Africa, and East Asia, for instance, have relatively comparable estimated cultural intolerance but significantly different dynamics of integration due to a substantial heterogeneity in their rates of homogamy and fertility. Indeed, our empirical analysis clearly shows that both selection into homogamous marriages and selection through fertility act as socialization mechanisms, affecting the dynamics of integration of cultural-ethnic minorities.

To examine in depth the mechanisms driving integration at equilibrium, we proceed by counterfactual analysis, connecting outcomes to reduced form results from the literature on immigrants' assimilation. More specifically, we study how integration responds to variations both in the supply of acceptance of the immigrants' cultural diversity on the part of natives, as well as in the demand of immigrants to preserve their cultural identity. Our first counterfactual simulation entails setting the cultural intolerance of natives with respect to all ethnic minorities equal to zero, so that natives offer complete acceptance of the immigrants' cultural diversity. In this case, we observe a significantly lower integration of second-generation immigrants: the fraction of the population composed of non-integrated immigrants decrease by 25% over the period of a generation. Facing no bias in the marriage market with natives, nor any preferences for culturally integrated children on the part of the native spouse in intermarriages, immigrants can achieve higher socialization rates when married with natives. At equilibrium, heterogamy increases, and so does socialization and fertility in intermarriages, reducing integration overall. In our second counterfactual simulation, we strengthen the dominance of cultural identity in the demand of immigrants, which could be due to a reduction in their economic incentives to integration. Interestingly, in this case, we find that cultural convergence is accelerated in a period of a generation, by 10 percentage points compared to baseline for a 20% increase in cultural intolerances. This result is the outcome of a lower participation in the marriage market and lower fertility of immigrants, motivated by the fact that the stronger attachment to their identities makes marriage a riskier and

costlier investment. This is the case for both homogamous marriages of immigrants and heterogamous marriages. In other words, in this counterfactual, the acceleration of integration is mostly an effect of the reduction in the population growth for immigrants with respect to natives. The probability that a child with an immigrant parent is integrated to the Italian culture is lower in the counterfactual with higher cultural intolerance, because immigrant parents' socialization is more effective, but the fraction of the total population (immigrants and natives) belonging to the native culture, on the contrary, is higher.

We also present a dynamic welfare analysis of our equilibrium model. We focus on a specific policy choice, i.e., a policy that strengthens the ethnic network of immigrants in society, fostering their ability to pass on their cultural identity to new generations with no direct investments from parents (e.g., public housing, freedom of religion, schooling). In our framework, this is captured indirectly by a parameter controlling the segregation of minorities. We compute a utilitarian social welfare function over multiple periods for different levels of segregation. From a dynamic perspective, our model displays an externality in households' fertility and socialization choices: individual households are myopic and do not internalize the effects of their current choices on the future size and composition of the population by cultural-ethnic traits. Due to the demand of minorities to preserve their cultural identity, we show that policies increasing segregation (e.g., by strengthening the relationships with co-ethnics) lead to positive contemporaneous and future welfare effects. Segregation policies redistribute contemporaneous welfare from agents in heterogamous marriages to both immigrants and natives in homogamous marriages so that both natives and immigrants overall realize welfare gains. The most striking effects appear though, in the future, when the externality in the fertility and socialization choices of households plays a fundamental role. A larger segregation of minorities in the present induces a larger number of second-generation immigrants in the marriage market in the future, and hence an increase in homogamous marriages of immigrants and in heterogamous marriages of immigrants with natives. Both contribute positively to the social welfare and more than compensate for the reduction in homogamous marriages of natives (which have constant value but decrease in numbers).

Finally, we take advantage of our model of cultural integration to investigate the effects of a surge in migration inflows on cultural heterogeneity. Doubling the number of second-generation minorities (keeping population shares constant across minorities), the integration rate lowers to 87% by the third generation compared to the 97% at the baseline. The effects are heterogeneous across cultural-ethnic groups: the integration rate is reduced by 11 and 5 percentage points, respectively, for immigrants from Sub-Saharan Africa and East Asia.

After a discussion of the related literature in the next section, the paper is organized as follows. Section 2 introduces an empirical analysis of marriage, fertility and cultural-

ethnic socialization of children by cultural-ethnic group. Section 3 outlines our theoretical framework, and Section 4 presents the structural model, the estimation strategy, and the identification of model parameters. We present the estimation results and various counterfactual simulations in Section 5 and 6. The welfare analysis of segregation policies is in Section 7, and the simulations of migration inflows in Section 8. Finally, Section 9 concludes.

1.1 Related literature

Our paper combines insights from the literature on the intergenerational transmission of cultural traits with those of family economics studies on the estimation of marital preferences. Methodologically, we embed a collective household decision problem into a matching model, as first in Chiappori et al. (2017, 2018); other papers along these lines include Gayle and Shephard (2019) and Galichon et al. (2019).[6] Marital utilities emerge endogenously as a function of fertility and intra-household inter-generational socialization choices, along the lines of Bisin et al. (2004).[7] In this way, we account for the fact that fertility and child-rearing are two key motives behind marriage (Browning et al., 2014), and at the same time we investigate the mechanisms that make ethnic-cultural traits a crucial dimension of marital matching (Bisin et al., 2004; Ciscato and Weber, 2016).

In terms of research question, this paper fits into the large empirical literature on the cultural integration of immigrants. Several of these studies concentrate on the immigrants' demand to preserve their cultural identity, by exploring socialization via children's first names and home language (Abramitzky et al., 2020; Fouka, 2020), intermarriage patterns (Gordon, 1964; Meng and Gregory, 2005; Furtado and Trejo, 2013; Guirkinger et al., 2019), self-reported national identity (Manning and Roy, 2010), contraceptive usage by teenage females (Achard, 2020), and neighborhood sorting (Hwang, 2019). Relatedly, some papers study the effects of specific immigration policies and reforms (Fouka, 2020; Abdelgadir and Fouka, 2019); others document the salience of the cultural identity across immigrant groups living in the same host country on a wide variety of outcomes (Fernández and Fogli, 2006, 2009; Fernández, 2011; Giuliano, 2007; Alesina et al., 2013). Separately, a sizable literature has also focused on the supply side of immigrants' integration, studying the economic roots of anti-immigrant sentiments on the part of natives (as surveyed in Borjas, 2014; Card and

[6] See Choo and Siow (2006); Chiappori et al. (2009); Dupuy and Galichon (2014); Choo (2015); Ahn (2018); Ashraf et al. (2020); Corno et al. (2020) for the more recent contributions to the study of marital matching problems in different contexts, and Chiappori and Salanié (2016) and Chiappori (2020) for a comprehensive review; and see also Lundberg and Pollak (1993); Chiappori (1988, 1992); Chiappori et al. (2002); Del Boca et al. (2014); Voena (2015) for advances in the study of spouses interactions in marriage.

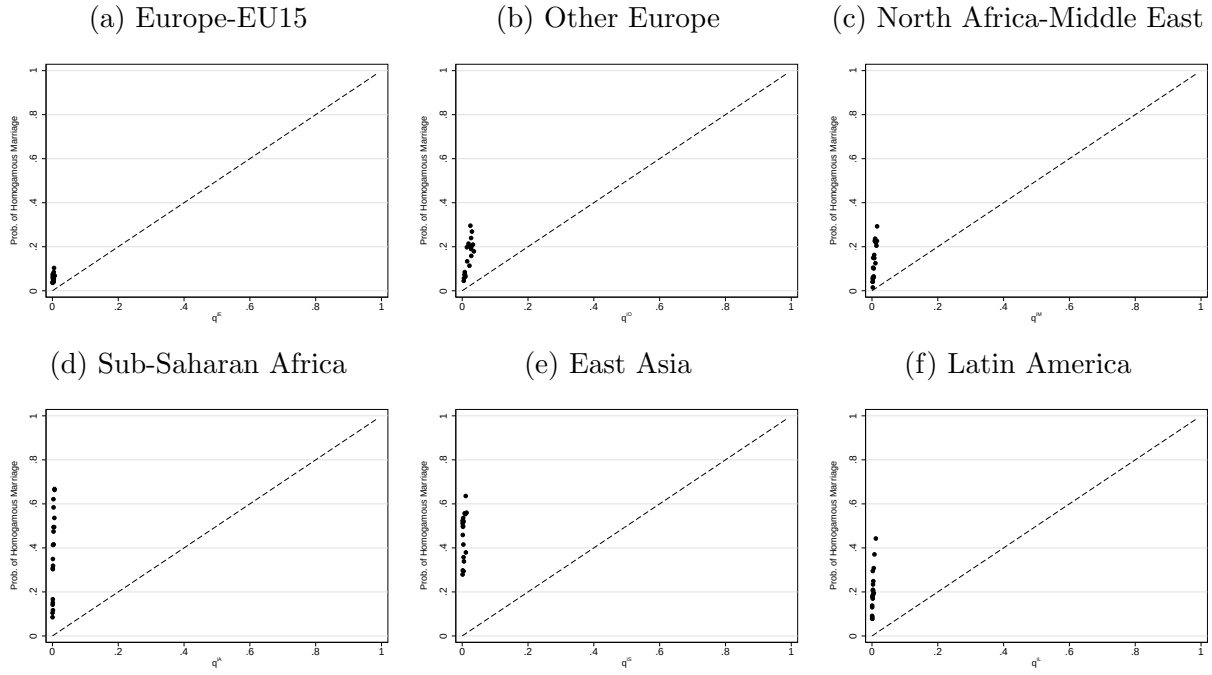
[7] See Bisin and Verdier (2000, 2001) for theoretical models of cultural transmission, and also Bisin and Verdier (2011) for a survey of the theoretical and empirical literature on the subject.

Peri, 2016; Dustmann et al., 2016), and their consequent political reactions (Dustmann et al., 2019; Tabellini, 2020). With respect to these studies, our structural approach allows us to identify and estimate both the demand of immigrants to preserve their cultural identity and the supply of acceptance of the immigrants' cultural diversity on the part of natives. Furthermore, our approach allows us to identify the fundamental interactions between the cultural trait and the education of the parents in their preferences over spouses and over the culture and education of the children. Within this unified and coherent framework, we can then study how the dynamics of immigrants' integration over time respond in equilibrium to variations in the preferences of both immigrants and natives, and we can investigate the effects of counterfactual policies.

2 Marriage by cultural group and education

In this section we introduce our empirical analysis of marriage by cultural-ethnic group and by education of the spouses, illustrating interesting stylized patterns. We first document the main marriage patterns in the data. We then link the observed marriage patterns, by cul-

Figure 1: Frequencies of Homogamous Marriage by Ethnic Group of Minorities



Notes: This Figure shows the probability that a member of a specific cultural-ethnic minority marries homogamously, in comparison with the probability of random matching (corresponding to the 45-degree line) implied by the distribution of cultural-ethnic traits across regions (reported on the horizontal axis). The distribution of q^i is averaged over the period 1995-2012.

ture and education, with several relevant marriage outcomes; specifically fertility, separation rates, children's cultural socialization, and children's education. While we conclude that cultural and educational attributes of spouses play a primary role in the marriage market, this analysis is necessarily silent about the identification of the fundamental mechanisms underlying immigrants integration outcomes. Indeed, we interpret this section as providing a striking motivation for the structural analysis of marriage we pursue in the rest of the paper.

The analysis takes advantage of administrative individual-level data from the Italian Statistical Institute (ISTAT, ADELE Laboratory), covering the universe of marriages celebrated in Italy from 1995 to 2012 (over 4 million marriages). In this section we briefly present the data, while referring to Appendix A for a comprehensive discussion of the sources.

In terms of cultural-ethnic traits, we distinguish between Italians, the native majoritarian group, and six immigrant minorities, aggregated by country of origin: European (EU15 countries), Other European, [8] North African and Middle-Eastern, Sub-Saharan African, East Asian, and Latin American; Appendix Figure C.1 contains a detailed analysis of this classification. Regarding education, we distinguish two groups: individuals with high (more than high school) and with low (no high school) education (we refer to these groups also as skilled and unskilled). Geographically, we divide Italy into 20 distinct marriage markets, corresponding to the administrative regions in the country.

2.1 Marriage

In our sample, 87.6% of marital unions are homogamous marriages of natives, while heterogamous marriages between a native and an immigrant spouse amount to 10.5%. Appendix Table C.2 reports the aggregate marriage distribution by ethnic group and education of spouses. Figure 1 reports the homogamy rates for each cultural-ethnic minority, i.e., the fraction of the members of a specific cultural-ethnic group that marries within their group, by geographical region. It documents high homogamy rates for all cultural minorities, significantly higher than those implied by random matching; corresponding to the 45-degree line in the Figure.

Beyond cultural origins, individuals also match along other observable characteristics. In particular, a large literature has consistently documented positive assortative mating by education (among many others, Becker, 1973, 1974; Dupuy and Galichon, 2014; Siow, 2015; Chiappori et al., 2017; Low, 2024). Our data is no exception along these lines. Indeed, in

[8] This group includes Eastern European countries which became EU members after the enlargements in 2004 and 2007. These enlargements altered the incentives of some immigrants to marry natives, hence the composition of intermarriages in the data is an average over the period; see Adda et al. (2020).

46 percent of marriages in our sample, both spouses are high-educated and 31 percent are both low educated; see Appendix Table C.2.

The homogamy and heterogamy rates clearly suggest that both cultural and educational factors constitute an important determinant of marriage choices. However, these rates depend on the distribution of men and women by cultural group and education in the marriage market under consideration. To control for both the uneven distribution of cultural traits and education (skill composition) in the population and potential gender imbalances, we turn to study *gains from marriage*, i.e., a measure of the utility gains associated to a marriage, with respect to the outside option of remaining single, for each marriage type (Choo and Siow, 2006).[9] Gains are computed by scaling the number of marriages for each marriage type by the geometric average of the number of singles of those types. Appendix Table C.3 reports average gains from marriage when marriage type is defined by the cultural traits and the educational characteristics of the spouses.[10] It shows that, as suggested by marriage rates, homogamy - across both cultural and educational lines - is strongly associated with higher gains from marriage.

As a first attempt at disentangling cultural and educational factors in marriage, Table 1 reports the results of regressing gains from marriage on marital sorting by culture and education. We document the presence of strong positive assortative matching by cultural-ethnic lines (column 1), as well as positive sorting by education (column 2). Importantly, the strength of sorting by culture is twice as high as the strength of sorting by education, even accounting for systematic differences in spouses' educational sorting when we distinguish between homogamous couples of both high-educated and low-educated spouses (columns 3 and 4).[11] Finally, Table 1 documents a positive and statistically significant interaction

[9]Since Choo and Siow (2006), gains from marriage have been studied in different contexts of interest; see references in Footnote 6. Gains from marriage are defined with respect to a predetermined set of marriage types; e.g., cultural and educational characteristics of the spouses. Precisely, consider a large marriage market, with a population of men and women, heterogeneous in terms of some relevant list of characteristics. Let μ_{hj} denote the fraction of marriages in the economy between a man with characteristics h and a woman j; let $\mu_{h.}$ and $\mu_{.j}$ denote the fraction of single men of characteristics h and single women of characteristic j, respectively. Gains from marriage G_{hj} are the utility gains associated to a hj marriage type, with respect to the case in which both spouses remain single, for each hj marriage in each market. Specifically, under Choo and Siow (2006)'s assumptions, gains are point identified from the distribution of marriages by type and the distribution of traits in each marriage market. We estimate gains from marriage, G_{hj}, by cultural-ethnic group and education of spouses hj, as follows:

$$G_{hj} = \log \frac{(\mu_{hj})^2}{\mu_{h.} \cdot \mu_{.j}}. \tag{1}$$

[10]To this end, we complement marriage data with population Census data in 2001 and 2011 to obtain information about single men and women.

[11]In Appendix Table C.4 we report on a further empirical exercise, which turns out to strengthen this

Table 1: Gains from Marriage by Culture and Education

Dep. var:	(1)	(2)	(3)	(4)	(5)
	Gains from Marriage				
Homogamous by culture	4.772***			4.806***	4.438***
	(0.205)			(0.205)	(0.211)
Homogamous by education		2.081***			
		(0.046)			
High-high educ			2.000***	2.042***	1.948***
			(0.062)	(0.065)	(0.071)
Low-low educ			2.163***	2.216***	1.999***
			(0.091)	(0.089)	(0.088)
Hom by culture × High-high educ					0.439***
					(0.095)
Hom by culture × Low-low educ					1.016***
					(0.091)
Observations	2456	2456	2456	2456	2456
R-squared	0.630	0.361	0.361	0.752	0.756
Marriage market (region) FE	Yes	Yes	Yes	Yes	Yes
Ethnic and Education (wife) FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable in the regression is gains from marriage, computed as in equation (1), across cells defined by region, spouses' cultural ethnic-group and education. The explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and education (*Homogamous by education*). Columns 3-5 include dummies for homogamous marriages by education, dividing marriages of high-educated (*High-high educ*) and low-educated (*Low-low educ*) spouses. Column 5 includes interaction variables. All specifications include marriage market (at the regional level) FEs effects, as well as wife educational level and cultural-ethnic group of origin FEs. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

between cultural and educational homogamy. Interestingly, homogamous cultural marriages of unskilled spouses exhibit higher gains from marriage compared to similar homogamous cultural marriages of skilled spouses, suggesting stronger complementarities in marriage.

2.2 Fertility and children investments

To further unpack the role of culture and education on marriage gains, we now study the mechanisms through which cultural and educational characteristics of the spouses associate

result, suggesting that cultural differences affect marriage allocation beyond educational incentives. We show that gains from marriage are negatively correlated to various measures of cultural distance of the ethnic traits of the spouses, e.g., genetic distance, language distance, and others, commonly used in the literature (Spolaore and Wacziarg, 2009, 2016). This correlation is statistically significant, consistently for different measures of cultural distance, even after conditioning on education.

with the observed marriage pattern in the data and hence with marriage gains.

First of all, cultural and educational characteristics may be associated with marriage gains *per se*, e.g., through the emotional and economic relationship between spouses. However, cultural and educational characteristics may be associated with marriage gains also through various other expected outcomes of marriages. In this respect, we systematically describe in this section the relationship between parental cultural homogamy and parental education with fertility, the separation rate, as a measure of lack of success in marriage, and with varios characteristics of the children which parents might have altruistic preferences for, notably cultural socialization and education. We contend, in particular, that marriage gains might be related to the expected quantity-quality dimension of children and that this dimension might be modulated by parental attitudes and preferences towards culture and education. Parents, for instance, might have a preference for children sharing their own cultural-ethnic traits and for educated children.

Consider fertility and separation rates first. Fertility rates by type of marriage come from municipality birth registries, which contain the universe of individual birth records of residents in the municipality of enrolment. In the sample, the fertility rate corresponds to 69.6% with an average of 1.54 children per family. Separation data, instead, come from the registries of civil court chancelleries and cover the universe of legal separations registered in Italy; see Appendix A for further details. Of all marriages, 7% end up in separation in the first years of the marital union.

Focusing on fertility, Table 2 documents that culturally homogamous marriages display systematically higher fertility, independently of the spouses' education, by 0.28 percentage points, on average, on a baseline fertility rate of 0.67 for heterogamous marriages. The effect of education on fertility, instead, is not statistically significant on average, nor focusing on marriages of spouses with similar education (column 2), nor focusing on marriages with at least one educated spouse (columns 3 and 4). In column 5, we document that the effect of education on fertility is mediated by cultural homogamy, i.e., it is positive for culturally heterogamous marriages and negative for culturally homogamous marriages. Specifically, among culturally homogamous marriages the fertility rate is higher for marriages of low educated spouses compared to marriages with at least one educated spouse, while the pattern is reversed for culturally heterogamous couples. Results are robust to different classifications of sorting by education; see Appendix Table C.5.[12]

This differential role of education on fertility - mediated by the cultural characteristic

[12]For a graphical representation, see Appendix Figure C.2. Interestingly, fertility displays significant variation across cultural-ethnic groups. This heterogeneity is at least partly explained by cultural differences; Panel B of Appendix Table C.4 shows that also fertility is negatively correlated to different measures cultural of distance, even after controlling for spouses' education.

of the marriage - appears to be a systematic feature of the data. Indeed, this is the case also for the separation rate. Appendix Table C.6 documents that culturally homogamous couples separate less, while the correlation of education with the separation rate depends on the cultural characteristics of the marriage; that is, it is negative for culturally heterogamous couples and positive for homogamous couples.

Turning to children's socialization and education, we adopt the same education classification for the children as for the parents and we adopt "speaking Italian at home" (by children and young adults living with their parents) as a proxy for unsuccessful cultural socialization. These data are obtained from the *Condition and Social Integration of Foreign Nationals Survey*, conducted in 2011 and 2012 in all Italian regions on a sample of 9,600 families; see Appendix A for a detailed discussion of language socialization as a proxy of cultural socialization.[13] Table 3 reports the effects of parents' culture and education on children's socialization. Not surprisingly, cultural homogamy is strongly and positively related to socialization (negatively to "speaking Italian at home"). However, the role of parental

Table 2: Fertility Rate by Culture and Education

	(1)	(2)	(3)	(4)	(5)
Dep. var:			Fertility rate		
Homogamous by culture	0.280***		0.280***	0.280***	0.337***
	(0.016)		(0.017)	(0.017)	(0.023)
Homogamous by education		-0.010	-0.005		
		(0.009)	(0.009)		
High educ (at least one)				0.019	0.035**
				(0.013)	(0.015)
Hom by culture × High educ (at least one)					-0.076***
					(0.015)
Observations	2456	2456	2456	2456	2456
R-squared	0.296	0.105	0.315	0.315	0.317
Marriage market (region) FE	Yes	Yes	Yes	Yes	Yes
Spouses' Cultural-ethnic group FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable in the regression is the fertility rate in marriages defined by spouses' cultural ethnic-group and education, and by region. The explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and by education (*Homogamous by education*). We also include a dummy for marriages with at least one high educated spouses (*High educ, at least one*) and its interaction with the dummy for cultural homogamous marriages (*Hom by culture × High educ*). All specifications include marriage market (at the regional level) fixed effects as well as spouses' cultural-ethnic group of origin fixed effects, and a dummy for wife's high education. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

[13]To support our interpretation, in Section 5.3.2, we provide consistent evidence that speaking Italian at home is, in fact, associated with higher achievement and educational choices of immigrant students, weaker ethnic identity, and stronger attitudes towards social integration.

education is once again mediated by whether the marriage is homogamous or heterogamous by culture: high parental education has a significant negative role on socialization in homogamous marriages, increasing the frequency of speaking Italian at home. Again, results are robust to different classifications of sorting by education; see Appendix Table C.7.[14]

Finally, we study children's education. Not surprisingly, Appendix Table C.8 documents that the probability that a child is highly educated is positively correlated with the education of the parents. Interestingly, this correlation is stronger in culturally heterogamous rather than in culturally homogamous marriages.

To summarize, the evidence presented suggests that marriage gains are associated with the cultural and educational characteristics of the spouses. Furthermore, the cultural and

Table 3: Fraction of "Italian spoken at home" by Culture and Education

	(1)	(2)	(3)	(4)	(5)
Dep. var:	"Italian spoken at home"				
Homogamous by culture	-0.320***		-0.319***	-0.322***	-0.407***
	(0.023)		(0.022)	(0.023)	(0.029)
Homogamous by education		-0.005	-0.024		
		(0.027)	(0.022)		
High educ (at least one)				0.051	-0.022
				(0.057)	(0.056)
Hom by culture x High educ (at least one)					0.112***
					(0.030)
Observations	615	615	615	615	615
R-squared	0.604	0.085	0.608	0.608	0.612
Marriage market (region) FE	Yes	Yes	Yes	Yes	Yes
Spouses' Cultural-ethnic group FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable in the regression is the fraction of "Italian spoken at home" in marriages defined by spouses' cultural ethnic-group and education, and by region. The explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and by education (*Homogamous by education*). We also include a dummy for marriages with at least one educated spouses (*High educ*) and its interaction with the dummy for homogamous marriages by culture dummies for couples of both high educated spouses, and a dummy for mix educated marriages (high-low) and their interaction with homogamous couples by culture (*Hom by culture x High educ*). All specifications include marriage market (at the regional level) fixed effects as well as spouses' cultural-ethnic group of origin fixed effects, and a a dummy for wife's high education. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

[14]The probability that an immigrant parent speaks Italian with his/her child in homogamous marriages is generally half as high as in heterogamous marriages with another immigrant, on average; this is consistent with other evidence, using different measures of socialization (among others Dohmen et al., 2012; Fouka, 2020). However, the variation across cultural groups is high; see Figure C.3. Interestingly, the relationship between socialization rates and cultural distance in heterogamous marriages, reported in Panel C of Table C.4, has an opposite sign, though not significant: the larger the cultural distance the lower the socialization to native Italian language.

educational characteristics of the spouses are in turn systematically associated with various important outcomes of marriages, their fertility and stability, and various dimensions of children's quality.[15] The evidence also indicates that the effect of parental education on the various marriage outcomes we study is systematically mediated by the cultural characteristic of the marriage. The analysis is however silent about the underlying possible interactions between preferences along cultural and educational lines. Indeed, it does not uncover any systematic relationship between parents' culture and education and their ordering of the characteristics contributing to children's quality as well as their representation of the quantity-quality trade-off. This is the manifestation of a a severe limitation of any reduced form analysis of marriage gains, necessarily falling short of identifying the fundamental aspects of individual preferences as they are manifested in the marriage market; and hence of identifying the determinants of the observed cultural integration pattern of immigrants.

Most specifically, the analysis of marriage gains we pursued in this section cannot separate preferences ranking spouses by education level from preferences valuing educational homogamy *per se* in marriage. For instance, marriage gains do not reveal whether highly educated individuals (in some cultural-ethnic groups) value cultural identify less - and hence, e.g., they are less fertile and socialize their children less, in a quality-quantity trade-off logic. Furthermore, this analysis cannot allow for a decomposition of how much marriage gains are determined by the immigrants' preferences to preserve their cultural identity or by different forms of the natives' (lack of) acceptance of immigrants' cultural-ethnic traits - what we referred in the Introduction as the demand and the supply of integration.

Motivated by these limitations in the reduced form analysis of marriage gains, we turn to a structural model in which marriage gains represent the deep preferences of individuals with respect to the outcomes of marriages - their stability, fertility, children's socialization and children's education. We allow these preferences to depend on the cultural trait and the education level of the individual - and we allow these two dimensions to interact. We also allow spouses in marriage to choose optimally, given their preferences, both fertility and children socialization. We postulate instead a direct (exogenous) link between marriage type and children education, calibrating it from the data. We then structurally estimate this equilibrium model of marriage, fertility, and socialization. This exercise allows us to identify the distribution of cultural preferences along the cultural-ethnic dimension as well

[15]The result that children's socialization and education represent fundamental components of marriage gains - and hence as determinants of marriage allocations - is consistent with a large literature, in economics as well as in the social sciences more generally. On the role of cultural-ethnic socialization, see e.g., the fundamental work of Boas (1928); Lévi-Strauss (1949) in anthropology; see also Riesman (1992); Smith (1996); Mayer (2013), and, in economics, Bisin and Verdier (2000); Bisin et al. (2004). On the role of children's education see Becker (1973, 1974); Chiappori et al. (2017, 2018).

as any structural relationship between education and cultural preferences, which in turn induces a structural relationship between education, cultural preferences and fertility and socialization. Furthermore, very importantly, this structural analysis allows us to identify preferences by cultural group - therefore allowing us to capture preference heterogeneity by cultural-ethnic trait. Indeed, substantial heterogeneity across cultural-ethnic groups in marriage patterns is one of the fundamental characteristics of our data. This is the case, e.g., with respect to i) homogamy, which is e.g., particularly high for Sub-Saharan African and East Asian minorities; ii) selection into marriage, whereby the probability of marrying is e.g., as high as 63% for Latin American individuals and only 46% for North African-Middle Eastern individuals.

3 A model of marriage and intra-family choices

Consider a frictionless marriage market. Individuals are heterogeneous in terms of their characteristics. These characteristics, for both males and females, consist of the cultural-ethnic trait and the education level. The set of characteristics is denoted C. For simplicity, we present the theoretical model for dichotomous cultural traits c: say $c = n$ for natives and $c = i$ for immigrants, and educational levels e: say $e = s$ for high educated (skilled) and $e = u$ for low educated (unskilled). The set of characteristics is then $C = \{ns, nu, is, iu\}$. In the empirical analysis, we shall allow for multiple cultural-ethnic traits of immigrants, while we remain with dichotomous education.

Individuals match in marriage anticipating the utility of their future choices as a household. Utility is transferable (TU) across spouses; that is, the utility possibility frontier is linear.[16] Transfers are endogenously determined as equilibrium outcomes, depending on the quality of the specific match but also on the set of available opportunities in the marriage market. Households' heterogeneity builds on the individuals'. The notation $\{h, j\}$ denotes a household type where the male has characteristic h and the female j; let $\{h, .\}$ denote the household type composed of a single male with characteristic h and $\{., j\}$ the one composed of a single female with characteristic j. T denotes instead the set of possible types of household, including those composed of single individuals. Abusing notation, we use $t \in T$ to index all types of households, and $hj \in C \times C \subset T$ to index married households.

Total marital utility is the sum of two components: i) a systematic component, depending on the marriage outcomes and intra-household choices (in turn depending on the spouses' future choices), and ii) an idiosyncratic component, capturing residual idiosyncratic returns

[16] Under TU, household decisions about public goods are independent of the structure of the marriage market and of the allocation of power between spouses; see Chiappori et al. (2015) for a discussion.

from marriage, observed by the individuals before marriage. Let ϵ_h and η_j denote the individual idiosyncratic preference shocks for men and women, respectively, with identity h and j. Both ϵ_h and η_j are vectors, each element of which represents the idiosyncratic component of utility associated with a possible type of spouse the individual might be matched with (including none, if he/she stayed single). The total expected utility of a household of type hj between man m with identity h and female f of identity j is assumed additive and separable in the shocks, $U_{hj} + \epsilon_{hj} + \eta_{jh}$ (resp. $U_{h.} + \epsilon_{h.}$). In the marriage market, individuals observe their idiosyncratic shocks and match along cultural-ethnic traits and educational levels, anticipating their marital utility $U_t + \epsilon_t + \eta_t$ for all different potential matches.

Let μ_t denote the fraction of marriages of type t formed in the population. Let m_h and f_j denote males and females with trait h and j, in turn, in the marriage market. Under separability and appropriate distributional assumptions on the individual unobserved heterogeneity components, the optimal stable assignment, μ_t, is the solution of the following convex problem, subject to the feasibility constraints:

$$\begin{array}{ll} \max_{(\mu_t \geq 0)_{t \in T}} & \sum_{t \in T} \mu_t U_t - \varepsilon(\mu) \\ s.t. & \\ \sum_j \mu_{hj} + \mu_{h.} = m_h & \forall h \in C, \\ \sum_h \mu_{hj} + \mu_{.j} = f_j & \forall j \in C \end{array} \tag{2}$$

where $\varepsilon(\mu)$ represents the generalized entropy of the matching, which captures the dispersion of individual preferences with respect to the aggregate preferences, conditional on spouses' attributes (Galichon and Salanié, 2017, 2021). Following Choo and Siow (2006), we assume that ϵ_t, η_t are independent and identically distributed random variable vectors with a type I extreme-value distribution (Gumbel). The matching model in (2) translates into a multinomial logit model (McFadden, 1974).[17] We further postulate that the systematic utility of marriage U_{hj} has two components:

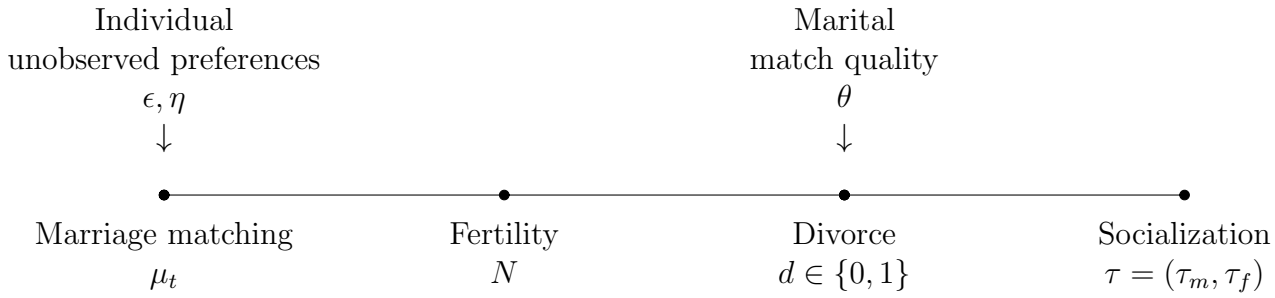
$$U_{hj} = U_{hj}^{par} + U_{hj}^{ec}.$$

The first, U_{hj}^{par}, is the utility deriving from the spouses' parental activities (single individuals

[17] Under this separability assumption, Galichon and Salanié (2021) and Chiappori et al. (2017) specifically show that the two-sided matching problem reduces to a series of one-sided discrete choice problems. From Galichon and Salanié (2021), the generalized entropy in (2) is:

$$\varepsilon(\mu) = \sum_{\substack{h \in C \\ j \in \{C,.\}}} \mu_{hj} \log \mu_{j|h} + \sum_{\substack{h \in \{C,.\} \\ j \in C}} \mu_{hj} \log \mu_{h|j}. \tag{3}$$

Figure 2: Timing of the Model



have no parental utility, $U^{par}_{h.} = U^{par}_{.j} = 0$). The parental utility component, U^{par}_{hj}, for each household type $hj \in T$, is the indirect utility of the spouses' future choices of fertility, socialization, and divorce which we obtain from the structural model we construct next. The second component, U^{ec}_{hj}, is the residual utility deriving from the spouses' economic activities, e.g., from their (present discounted) income. The economic component of the hj-type marriage utility is assumed proportional to the sum of the utility of an h-type man and a j-type women, were they stayed single:

$$U^{ec}_{hj} = \alpha \left(U_{h.} + U_{.j} \right);$$

where the parameter α captures the relative effects of marriage on the economic opportunities for the spouses.

The timing of the model is illustrated in Figure 2. After households are formed in the marriage market, in the second stage, the spouses in the household choose, cooperatively, fertility; that is, the number of children, N. In the third stage, a match (not individual) specific quality shock θ is realized, which is observed by the spouses. Depending to the realization of the shock, the spouses cooperatively decide $d(\theta)$; that is, whether they remain married with $d = 0$, or to divorce with $d = 1$.

Finally, parents invest in the cultural socialization of children. They exert costly effort in order to shape their children's cultural indentity. More specifically, parents choose a cultural socialization effort, $\tau(d)$, either as a cooperative decision by both parents in the household or as a non-cooperative decision of the mother in case the household is separated.[18] Socialization effort of a parent of cultural-ethnic c increases the probability that children will share the

[18]The fact that fertility is chosen before the realization of θ (hence constant within household type) and that socialization effort is chosen after the divorce decision are mere simplifications. What is important for our analysis is that, on the one hand, fertility varies systematically across household types and that the divorce decision affects the socialization effort.

same trait. When parents choose socialization they account for the probability that children will be high or low educated, understanding that this probability will depend on their own (both parents) education level. In this respect we dramatically simplify the analysis by not modeling explicitly a choice of parents with respect to their children education, but assume they rationally anticipate the probabilities in the data.

We consider a series of other simplifications in the model. First of all, the marital utility is assumed proportional to fertility N and investment in fertility entails a cost $\kappa(N)$, increasing and convex in N. Secondly, each child is associated to cultural-ethnic trait and an educational level, a characteristic $k \in C$, representing his/her quality from the point of view of the parents. Parental utility per child is then composed of i) an utility component deriving from the child's characteristic, u_{hj}^k; and ii) a direct utility component δ deriving from having a child in the marriage as opposed to outside the marriage, that is, the spouses choose to separate, $d = 1$. Letting $u_{hj} = E(u_{hj}^k)$ and assuming that the marital quality shock θ enters marital utility only if the household does not divorce - that is, if it chooses $d = 0$, we have:

$$U_{hj}^{par} = N\left[u_{hj} + \delta(1 - d(\theta))\right] + E\left(\theta(1 - d(\theta))\right) - \kappa(N) \tag{4}$$

From the above equation notice that the systematic value of a match is a function of u_{hj}, the (match specific) expected utility that parents derive from children's characteristics. This expected utility is endogenous in that the cultural dimension of children's quality characteristics depends on parental socialization choice, which we turn to introduce.

Let P_{hj}^k denote the probability that a household $hj \in T$ has a child $k \in C$. We assume the following form independence: $P_{hj}^k = P_{hj}^c P_{hj}^e$, where $c \in \{n, i\}$ and $e \in \{s, u\}$. Now P_{hj}^e is exogenous and directly observed in the data. P_{hj}^c is endogenous and we shall construct it next. Let V_h^k denote the utility of a parent of type $h \in C$ for a child of type $k \in C$, for $C = \{ns, nu, is, iu\}$. We can then write

$$u_{hj} = P_{hj}^s \left[\sum_{c \in \{n,i\}} P_{hj}^c \left(V_h^{cs} + V_j^{cs}\right) \right] + P_{hj}^u \left[\sum_{c \in \{n,i\}} P_{hj}^c \left(V_h^{cu} + V_j^{cu}\right) \right]; \tag{5}$$

Parental preferences over the cultural traits and education of their children depend on the parents' culture and education. We restrict the form of such dependence as follows:

$V_h^{cs} = V_h^{cu} + S_c$, for any $c \in \{n, i\}$; so that S_c is the additive value of education (of a child $e = s$ as opposed to $e = u$) of parents in culture c;

$V_{cu}^{c'u} = \gamma_c V_{cs}^{c'u}$, for any $c, c' \in \{n, i\}$; so that γ_c is a multiplicative distinctive shifter of

the cultural preferences of educated parents in culture c.

Imposing then the normalization that $V_{cs}^{cs} = V$, for all groups $c \in \{n, i\}$, we obtain the following relationships between preferences over culture and education:

$$V_{cu}^{c'u} = \gamma_c V_{cs}^{c'u}; \quad V_{cu}^{c's} = \gamma_c V_{cs}^{c's}$$
$$V_{cu}^{c's} = V_{cu}^{c'u} + S_c; \quad V_{cs}^{c's} = V_{cs}^{c'u} + S_c;$$

and hence

$$V_{cu}^{c's} = \gamma_c V_{cs}^{c'u} + S_c; \quad V_{cu}^{c's} = \gamma_c (V_{cs}^{c's} - S_c) + S_c = \gamma_c V_{cs}^{c's} + (1 - \gamma_c) S_c.$$

We can now construct U_{hj}^{par} backwards, from socialization to divorce and, finally, to fertility.

Socialization. We start from the socialization problem, given (N, θ, d). In fact, under the preference structure we imposed, the socialization choice is independent of fertility N and it depends on θ only through d. We assume that, conditionally on his/her education, each parent's preferences over the cultural-ethnic identity of his/her children are biased towards his/her own trait, as a manifestation of paternalistic altruism:

$$V_{ce}^{ce'} > V_{ce}^{c'e'}, \text{ for all } c \neq c' \in \{n, i\} \text{ and all } e, e' \in \{s, u\}.^{19}$$

We refer to $\Delta V_{ce}^{c'e} = V_{ce}^{ce} - V_{ce}^{c'e}$ as the *cultural intolerance* of cultural-ethnic group c with respect to group c', for $c \neq c' \in C$.

With regards to the socialization technology, following Cavalli-Sforza and Feldman (1981) and Bisin and Verdier (2001, 2011), we interpret the process of transmission of cultural traits as the interaction of two forces: the *vertical* socialization of parents within the family, and the *horizontal* socialization of the society at large. As for the vertical socialization at the level of the family, we introduce several simplification assumptions. First of all, within a family all children identify to the same trait.[20] Secondly, homogamous native households socialize their children with probability 1; that is, children of native parents speak the native language. Thirdly, in a household of type hj - with $h = ce$ and $j = c'e'$ - the socialization effort of the father, τ_m, has the objective and the effect of increasing the probability that the children identify with his trait, c; similarly, the socialization effort of the mother, τ_f, has the objective

[19] This assumption is instrumental in deriving the implication of the model, but it is indirectly tested in the empirical exercise.

[20] In particular, we abstract from differences in socialization preferences regarding the gender and/or the birth order of children, and from socialization externalities driven by spillover effects across siblings.

and the effect of increasing the probability that the children identify with her trait, c'. Parents in heterogamous households, such that $c \neq c'$, face conflicting incentives in the socialization of children, while parents in homogamous households, with $c = c'$, benefit from coordinated incentives. Thus, the value of the marriage increases in the coordination of investments in children's cultural socialization. Finally, the socialization technology responds to the social environment. Let q^c define the fraction of individuals with trait c in the overall population. We assume that if a child fails to be socialized within the family, horizontal socialization occurs mimicking a role model selected at random from the population of reference, with probability q^c.

Let $P^c_{hj}(\tau, d)$ denote the probability that a child in a family of type hj - with $h = ce$ and $j = c'e'$ is socialized with the father's trait $c \in \{n, i\}$, when the socialization effort is $\tau = (\tau_m, \tau_f)$ and the divorce choice is d. We assume natives in homogamous marriages socialize their children with certainty, $P^n_{hj}(\tau, 0) = P^n_{hj}(\tau, 1) = 1$, if $c = c' = n$; hence, extending Bisin and Verdier (2000), socialization technologies are as follows:

$$\begin{array}{ll}
P^i_{hj}(\tau,0) = \tau_m + \tau_f + (1-\tau_m-\tau_f)q^i, & P^n_{hj}(\tau,0) = (1-\tau_m-\tau_f)(1-q^i) \text{ if } c = c' = i \\
P^i_{hj}(\tau,0) = \tau_m + (1-\tau_m-\tau_f)q^i, & P^n_{hj}(\tau,0) = \tau_f + (1-\tau_m-\tau_f)(1-q^i) \text{ if } c = i, c' = n \\
P^i_{hj}(\tau,0) = \tau_f + (1-\tau_m-\tau_f)q^i, & P^n_{hj}(\tau,0) = \tau_m + (1-\tau_m-\tau_f)(1-q^i) \text{ if } c = n, c' = i.
\end{array} \tag{6}$$

Assuming that the mother is given custody of children in divorce, socialization probabilities under divorce $P^k_{hj}(\tau, 1)$ are as the ones reported in (6), after imposing $\tau_m = 0$.

The expected utility u_{hj}, as defined by Equation 5, depends on the socialization probabilities in 6, and hence on both the choice of τ on the part of parents and on whether they divorced or not $d = 0, 1$. Socialization effort τ is then the solution the following maximization problem:

$$\max_{\tau \geq 0} \; u_{hj}(\tau, d) \tag{7}$$

Let the solution be denoted $\tau(d)$.[21] Notice that it depends only on $\Delta V^{c'e}_{ce}$ rather than on the utility levels $V^{ce}_{c'e}$. Moreover, at the solution, the parents' choice of socialization effort is also a function of q^i, i.e., of the proportion of immigrants of group i in the reference population.

Divorce. After observing the realization of the marriage quality shock θ, the spouses optimally choose whether to dissolve the marriage (divorce) or not, rationally anticipating their total utility from the socialization process. Given N, a type hj household divorces, choosing

[21] Whenever possible without confusion, we avoid to use the hj subscript in the notation.

$d(\theta) = 1$, if

$$N\left(u_{hj}(\tau(1),1)\right) > N\left(\delta + u_{hj}(\tau(0),0)\right) + \theta$$

Given $F(\theta)$ the cumulative distribution of θ, the probability of divorce of a type hj household with N children is

$$\pi\left(N\right) = F\left(Nu_{hj}(\tau(1),1) - Nu_{hj}(\tau(0),0) - N\delta\right)$$

Fertility. The quantity-quality trade-off that characterizes endogenous fertility choices (Becker, 1960) is captured in the model, as the optimal number of children is determined by the expected socialization quality per child, interacted with the effect of fertility itself on dissolution, and the marginal cost of raising them:

$$\max_{N} \; N\left(\pi\left(N\right)u_{hj}(\tau(1),1) + \Big(1-\pi\left(N\right)\Big)\Big(\delta + u_{hj}(\tau(0),0)\Big)\right) - \kappa(N) \tag{8}$$

3.1 Discussion

Two main modeling choices we have adopted are worth discussing. First of all, our analysis in this paper aims at understanding immigrants' cultural integration patterns. As a consequence, the model is centered on the parent's individual choices which more directly determine cultural integration: fertility, and socialization, as a mechanism for cultural transmission. To streamline the analysis, the transmission of education is then calibrated from the data rather than modeled as an individual choice problem.

Secondly, the structure of individual preferences we postulate allows for substantial variation across the individuals' cultural-ethnic group and education class; that is, across the observable individual characteristics. Of course, variation across unobservable characteristics, e.g., idiosyncratic preferences for fertility, education, and socialization, could in principle be important in understanding differential behavioral patterns. However, when these individual characteristics are unobservable to the spouse entering the marriage - education, and culture will tend to be used as proxies. In any case, the effects of unobservables will be captured in our analysis as long as they are idiosyncratic after conditioning on preferences for culture and especially education.[22]

[22]In other words, we submit that omitted variable bias in our analysis is bound to be limited once we

3.2 Results

We describe here informally the most important implications of the model in the previous section, for a culturally heterogeneous society in which group i is a minority, $q^i \in (0, 1/2)$.

Socialization. Parents make costly investments in order to socialize their children, both in culturally homogamous and heterogamous families. Socialization investments in homogamous families benefit from coordinated incentives. Conversely, socialization investment in heterogamous families depends on cultural intolerance asymmetries. In addition, homogamous families, when married, hold a more efficient socialization technology, compared to heterogamous ones. If they divorce, the socialization technology is the same independently of the type of household. As a consequence,

In culturally homogamous minority households ii, when the parents stay married, both parents socialize the children. If, instead, the household divorces, only the mother has custody and socializes the children, by assumption, and the investment in socialization is lower. In culturally heterogamous households ni and in, when the parents stay married, only the parent with higher cultural intolerance has a strictly positive socialization effort. If, instead, the household divorces, in this case as well, only the mother socializes the children. Heterogamous households, contrary to homogamous ones, invest more in socialization when divorced than when married.

For all household types, married or divorced, the probability of successful socialization to the trait desired by the parents (or parent) doing the investment is greater than the rate associated to horizontal socialization.[23] We turn now to study comparative statics,

In culturally homogamous minority households, whether parents divorce or stay married, both parents' socialization efforts are monotonically increasing in cultural intolerance and decreasing in the size of their cultural group, q^i. In culturally heterogamous households, the socializing parent's effort is monotonically increasing in his/her own cultural intolerance; if parents stay married, the socializing parent's effort is also decreasing in his/her spouse's cultural intolerance. It is also the case that the minority i socializes more than the majority n, ceteris paribus.[24]

Divorce. Consider an household with positive fertility, $N > 0$. As the systematic gains from

allow for variation by cultural-ethnic group and education class.

[23] Except in the knife-edge case of heterogeneous household with equal cultural intolerance preferences, as in this case parents do not socialize children.

[24] This is a property called *cultural substitution* in Bisin and Verdier (2001).

marriage derive from socialization and divorce leads to a generally less efficient socialization technology,

All types of household $hj \in T$ stay married if their marriage quality shock is positive, $\theta_{hj} \geq 0$; they divorce only if the quality shock is negative and large enough (in absolute value).

On the other hand, in culturally heterogamous households, mothers have an advantage in socialization after divorce,[25]

The divorce probability of culturally heterogamous families is higher compared to homogamous minority families, for the same realization of the stability shock, θ_{hj}, if the mother has higher cultural intolerance. If, instead, the father has higher cultural intolerance, the divorce probability of culturally heterogamous families is higher compared to homogamous families if and only if the father belongs to the cultural-ethnic minority i.

More generally, our model displays a quantity-quality trade-off in fertility, since quality is effectively represented by the associated efficiency of socialization,

The divorce probabilities, for both culturally homogamous and heterogamous families, are decreasing in the number of children.

Fertility. The fertility rates for all types of households are strictly positive. The main result is that,

The fertility rate in culturally homogamous families is larger than the fertility rate in heterogamous families.

Matching. The systematic component of the marital utilities exhibits a form of endogenous complementarity in socialization technologies. As a consequence,

The optimal allocation in the marriage market generates positive assortative mating along cultural-ethnic lines. Individuals optimally select into homogamous families. Deviations from positive sorting are the result of the presence of heterogeneity in individual unobserved preferences and of potential market asymmetries in the distribution of cultural-ethnic traits between the two sides of the market.

Education. We allow the model to depend on spouses' education via two parameters, γ_c and S_c. The parameter γ_c measures the relative salience of cultural preferences for unskilled

[25]Divorce choices for heterogamous families might be interpreted as a strategic deviation from marriage for mothers who have a preference to socialize children, and expect to have a higher probability of child custody attainment; see Dohmen et al. (2012) for evidence.

parents. Therefore, socialization effort - and thus the probability of socialization to their own culture c - as well as fertility rates of unskilled parents are increasing in γ_c. The parameter S_c, instead, represents the parents' premium in having a high skilled child (which happens with a probability calibrated in the data). Socialization effort is therefore unaffected by S_c, while fertility choices are decreasing in S_c, the higher S_c the higher the loss in children quality for unskilled children. This comparative result holds true for all types of families, both homogamous and heterogamous, and independently from parents' education.

4 Structural estimation: Methodology

We estimate the parameters of the model by observing the marital matching patterns, as well as the fertility, separation, and socialization rates. Taking the model in Section 3 to data, we keep considering its extension to Z cultural-ethnic groups: Italians, denoted n, and 6 immigrants groups i: i_E for Europe-EU15; i_O for Other Europe; i_M for North Africa-Middle East; i_A for Sub-Saharan Africa; i_S for East Asia; i_L for Latin America. Thus, $c \in \{n, i\}$ and $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$. We consider R separate marriage markets. In the following, we introduce relevant assumptions and functional form parametrization, we describe parameters of interest, we introduce our estimation procedure, and we discuss identification.

4.1 From the model to the data

Recall that from the model in Section 3, the hj-type marriage utility is $U_{hj} = U_{hj}^{par} + U_{hj}^{ec}$. The model specifies a functional form for U_{hj}^{par} which depends on the structure of ΔV_h^k, for all $h, k \in C$ and q^c. The economic component of the $t = hj$-type marriage utility satisfies $U_{hj}^{ec} = \alpha_z \left(U_{h.} + U_{.j}\right)$.

Concerning socialization probabilities, we assume that in households of cultural type ii, in and ni children can only be socialized either to trait i or n. In a heterogamous household with both immigrants parents, on the other hand, the children can be socialized either to one of the parents' traits or to n.The remaining socialization probabilities are constrained to be zero. Also, we allow socialization and fertility costs to capture systematic differences between homogamous and heterogamous couples, indexed by $z \in \{het, hom\}$, respectively:

$$c(\tau) = \sigma_{\tau_z} \left\{ \lambda_{\tau_z} \frac{1}{2} \tau^2 + (1 - \lambda_{\tau_z}) \left(e^{\frac{\tau}{1-\tau}} - 1 \right) \right\}; \; \kappa(N) = \sigma_{N_z} \left\{ \lambda_{N_z} (N)^{\xi_s} + (1 - \lambda_{N_z}) \left(e^{N^{\xi_z}} - 1 \right) \right\},$$

where $\xi_z \geq 1$, captures the dependence of fertility costs on childbearing decisions.[26]

[26]Our parametrization of socialization and fertility costs guarantee that they are increasing and weakly

For flexibility in the estimation, we allow the distribution of θ_{hj} to have a mean that depends on the household type hj. More specifically, we assume that θ_{hj} follows a generalized logistic distribution with location a_{hj} and scale parameter b. We normalize $b = 1$ and we set a_{hj} to match the dissolution probability of couples *without children* in the data for all hj; i.e., $a_{hj} : F(0; a_{hj}, b) = \hat{\pi}_{hj}(0)$.[27] This assumption allows us to capture systematic differences in separation rates across household ethnic groups without children; that is, independently from children socialization mechanism. Aggregate evidence is reported in Appendix Table C.9.

We allow the residual value of marriage, i.e., the value of staying single, to vary with ethnic group and education and to be distinct across homogamous and heterogamous marriages, $\omega_{c_z}\bar{\omega}_{e_z}$, for $c \in \{n, i\}$, $e \in \{s, u\}$, and $z \in \{hom, het\}$. This is to capture, indirectly, differential sorting in both observables and unobservables across cultural-ethnic groups, educational attainment, and between families.

Finally, we model the role played by the immigrants' cultural-ethnic network within the transmission process, relaxing the initial assumption of unbiased horizontal socialization frequencies.[28] A strong network fosters the ability of immigrants' communities to pass on their cultural identity to new generations with no direct investments from parents (e.g., public housing, freedom of religion, schooling). In our framework, this is captured indirectly by introducing a positive segregation bias, ρ, allowing each minority i to face a segregated socialization pool composed of a fraction Q^i of individuals of the same group i; where

$$Q^i = \rho q^i \quad \forall i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}.$$

The horizontal socialization of the majority group is rescaled to represent its complement. The parameter ρ represents the strength of the contribution of group i in the socialization of new generations of minorities to trait i with respect to its actual representation in the population q^i under random matching. The higher is ρ, the more effective is the horizontal socialization of the society at large.

convex functions in the parents socialization efforts and childbearing choices, respectively, and they satisfy regularity Inada conditions for interior solutions. While this specification of fertility cost rules out returns to scale, our estimates point to strictly convex cost functions, suggesting this restriction is not binding.

[27] Because of data limitations, we estimate the probability of dissolution of couples without children, $\hat{\pi}_{hj}(0)$, as the linear combination of a match-specific component, to capture heterogeneity in divorce rates across matches, and a regional specific component, to capture heterogeneity across regions.

[28] We derive the population distribution by cultural-ethnic group and region for the period 1995-2012, from municipality records on the foreign resident population. Population shares by ethnic group and region are calculated thanks to administrative data on the total resident population by region. The maps in Figure C.4, display the geographical variability in the ethnic groups' distribution across markets.

4.2 Parameters and Estimation

The main parameters of interest are the cultural intolerance parameters, ΔV_h^k, for all $h, k \in C$. For identification purposes, we normalize $V = 100$, so that cultural intolerances are measured in units corresponding to percentages of V. We are left with $Z(Z-1) = 42$ cultural intolerance parameters to estimate. We also estimate education parameters, γ_c and S_c, for all c. The other parameters to be estimated are: socialization and fertility cost function parameters, $\sigma_{\tau_z}, \lambda_{\tau_z}$ and $\sigma_{N_z}, \lambda_{N_z}$; dependence of fertility costs on childbearing decisions, ξ_s; direct value of fertility (independently from cultural socialization), δ; segregation bias ρ; outside option of being single ω_{c_z} and $\bar{\omega}_{e_z}$ for all c, e and z; and relative effects of marriage on spouses' economic opportunities, α_z.

Let $\boldsymbol{\beta}$ denote the vector of parameters. Given an exogenous population distribution q^c, for all cultural groups c, and the exogenous probability of P_{hj}^e for all $hj \in T$, the structural model provides us with the theoretical moments in reduced form, $\widetilde{\boldsymbol{\Pi}}(\boldsymbol{\beta})$. Specifically, in our estimation, the theoretical moments we exploit are maps from $\boldsymbol{\beta}$ into U_{hj}, N_{hj}, π_{hj}, for all hj, and $P_{hj}^k(d)$ for all hj and k, and marital status d.[29]

The empirical moments are $\widehat{\boldsymbol{\Pi}} = \{\hat{U}_{hj}, \hat{N}_{hj}, \hat{\pi}_{hj}, \hat{P}_{hj}^k\}$, for all hj and k. In particular, we compute the implied marital surplus $\hat{U}_{hj}$ through the identification equation of the marital matching function in (1), where $\hat{\mu}_{hj}$ is obtained from the distribution of marriages over the period 1995-2012, while $\hat{\mu}_{h.}$ and $\hat{\mu}_{.j}$ are taken from the population vectors by ethnic group, education, gender and marital status of individual Census data in 2001 and 2011. We compute fertility rates $\hat{N}_{hj}$ as the average number of children in households of type hj, including zeros.[30] We evaluate separation rates $\hat{\pi}_{hj}$ as the fraction of marriages of type hj ending in separation during the period of analysis, conditional on having children. Finally, we construct socialization frequencies, $\hat{P}_{hj}^k$, as the fraction of households of type hj in which children speak a given language k at home.[31] Given normalization restrictions, we end up with a total of 101 parameters to match 2,416 moments.[32]

We estimate model parameters via a method of moments estimator, by matching the vec-

[29] Theoretical socialization moments are computed as follow. For given values of the parameters $\boldsymbol{\beta}$ and an exogenous population distribution q^c, for all c, first order conditions of the optimization problem in (7) pin down the optimal socialization effort $\tilde{\tau}(d)$, by means of cost function parametrization $c(\tau)$ in (4.1). Given optimal effort at the household level, we can compute the socialization frequencies implied by the model $P_{hj}^k(d)$ for all hj and k, and marital status d.

[30] We check that our results are robust to computing fertility moments differently, see Section 5.

[31] Because within each family socialization frequencies sum up to one, we exclude from the estimation redundant moments. Moreover, we exclude socialization moments for divorced families for data limitations.

[32] Since q^c is indexed by the region $r \in R$, we obtain a set of moments for each region r. We have hidden the index r in the dimensionality of the vectors of moments. See Appendix A for a detailed description of the empirical moments.

tor of theoretical moments implied by the model, $\widetilde{\mathbf{\Pi}}(\boldsymbol{\beta})$, for a specified choice of parameters $\boldsymbol{\beta}$, with their empirical counterparts observed in the data, $\widehat{\mathbf{\Pi}}$. Formally, given a weighting matrix $\mathbf{\Omega}$,[33]

$$\hat{\boldsymbol{\beta}} = arg\min_{\boldsymbol{\beta}}[\widehat{\mathbf{\Pi}} - \widetilde{\mathbf{\Pi}}(\boldsymbol{\beta})]^{\intercal}\mathbf{\Omega}[\widehat{\mathbf{\Pi}} - \widetilde{\mathbf{\Pi}}(\boldsymbol{\beta})].$$

4.3 Identification

Our estimation procedure exploits two sources of cross-sectional variation: variation across family types (by cultural groups and education) as well as variation in the composition of the population across regions. Identification, thus, requires us to assume that all parameters are constant across the 20 geographical regions $r \in R$ and that each region corresponds to a separate local marriage market.[34] Furthermore, identification hinges also on i) the random variable θ_{hj} having the same distribution across households hj; ii) the segregation bias, ρ, the economic complementarity in marriage, α_z, as well as the socialization and fertility cost parameters $(\sigma_{\tau_z}, \lambda_{\tau_z}, \sigma_{N_z}, \lambda_{N_z}, \xi_z, \delta)$ being independent across ethnic groups c and educational level e. Independence of costs, in particular, implies that any difference in costs across groups and education would be attributed, in our estimates, to cultural intolerances. On the other hand, we can allow the outside options of being single, ω_{c_z} and $\bar{\omega}_{e_z}$, to vary by cultural-groups, by education, and be specific to each family type. These parameters are pinned down by the average probability of marrying for each ethnic and educational group and family type across regions.

Under these assumptions, no restrictions need be imposed on cultural intolerance parameters: geographical variation in population vectors allows us to separately identify the cultural intolerance of parents of type h with respect to children of type k, ΔV_h^k, for all $h, k \in C$.[35] In particular, and most importantly, we can identify the cultural intolerance parameters of minorities with respect to children integrated as natives, $\Delta V_{ie}^{ne'}$, separately from cultural intolerances of natives versus all minorities, $\Delta V_{ne'}^{ie'}$, for all education levels $e, e' \in \{s, u\}$ and all immigrant group i. This is, in fact, tantamount to identifying demand

[33] Because of the uneven distribution of marriages in our sample, the weighting matrix is constructed by balancing sample size considerations and representation. Hence, we assign the same weight to homogamous marriages of natives and to the rest of marriages; in turn, the rest of the marriages are weighted by their relative representation in the data. We solve the optimization problem via the Differential Evolution (DE) algorithm, a global optimization algorithm, first introduced by Storn and Price (1997), designed for non-convex and non-linear programming problems with potentially multiple local optima.

[34] We calculate that in more that 92% of our marriages, spouses shared the same region of residence before the marriage.

[35] We should note, however, that while we allow cultural intolerance and education parameters to be cultural-ethnic group specific, we need to assume they are constant across households within cultural-ethnic group. In addition, we need to assume that preference parameters are constants along gender lines.

and supply components of cultural integration as an equilibrium phenomenon. Such identification is possible in our setup because theoretical moments for homogamous immigrant families are a function of demand parameters only, and thus identify demand. Conditional on demand, theoretical moments of heterogamous marriages with natives pin down supply parameters. Moreover, given the structure of preferences presented in Section 3, we are able to identify cultural-group specific education parameters, γ_c and S_c, for all c. A detailed proof of identification for the model in Section 3 is provided in Appendix D.

5 Structural estimation: Results

In this section, we start by describing the fit of the model. We then introduce the parameter estimates. Finally, we present various validation exercises and we discuss the main assumptions of our empirical model, reporting on several robustness checks. As for the implication of our estimates with regards to the evolution of cultural traits in the long-run and the mechanisms driving integration at equilibrium, we develop them in the next section.

Table 4: Fit of the Model

a. Homogamous Families

	Italian Soc						Fertility					
	High		Mix		Low		High		Mix		Low	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Italian	1.00	1.00	1.00	1.00	1.00	1.00	1.09	1.10	1.12	1.11	1.17	1.13
EU15	0.49	0.47	0.49	0.47	-	-	0.58	1.00	0.72	1.01	0.77	1.03
Eu Other	0.42	0.37	0.44	0.34	0.34	0.33	0.58	0.49	0.64	0.61	0.68	0.71
Middle East	0.32	0.33	0.41	0.32	0.28	0.32	0.79	0.68	0.79	0.74	0.79	0.79
Sub-Africa	0.52	0.34	0.43	0.35	0.35	0.35	0.86	0.85	0.85	0.84	0.84	0.83
East Asia	0.39	0.36	0.21	0.35	0.14	0.34	0.98	0.50	1.09	0.60	1.09	0.75
Latin America	0.41	0.40	0.54	0.40	0.46	0.40	0.44	0.51	0.40	0.39	0.36	0.32

b. Heterogamous Families

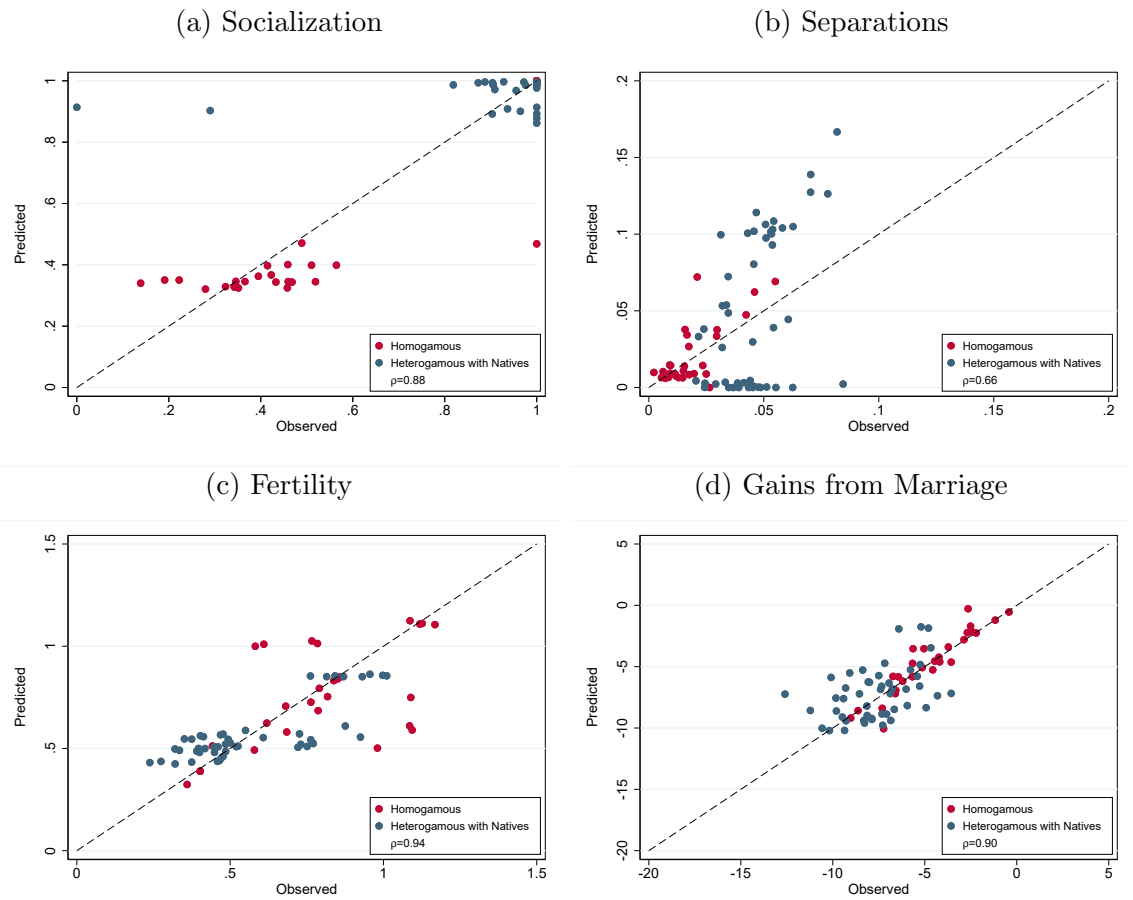
	Italian Soc						Fertility					
	High		Mix		Low		High		Mix		Low	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Italian	0.96	0.96	0.94	0.94	0.95	0.93	0.68	0.64	0.57	0.57	0.47	0.56
EU15	0.96	0.98	0.91	0.98	0.92	0.98	0.95	0.86	0.90	0.83	0.78	0.84
Eu Other	0.90	0.96	0.96	0.93	0.96	0.90	0.62	0.59	0.54	0.51	0.46	0.49
Middle East	1.00	0.94	0.95	0.93	0.99	0.98	0.48	0.39	0.42	0.44	0.28	0.40
Sub-Africa	1.00	0.98	0.92	0.98	0.87	0.98	0.60	0.54	0.51	0.52	0.37	0.53
East Asia	1.00	0.98	0.82	0.98	0.91	0.98	0.45	0.46	0.41	0.47	0.36	0.48
Latin America	1.00	0.87	0.91	0.89	0.94	0.91	0.53	0.55	0.45	0.52	0.35	0.48

Notes: This table shows the fit of the model by cultural-ethnic group of spouses, for homogamous (panel a) and heterogamous (panel b) families. Estimates are weighted by the number of marriages.

5.1 Model Fit

The model fits the data well. Table 4 compares the average observed and predicted socialization and fertility moments, for homogamous and heterogamous families, respectively dividing families by educational sorting. Similarly, the model fits well the fertility rates for homogamous and heterogamous families. In particular, the model is able to capture the differential pattern in fertility choices by educational sorting, that is, while high skilled homogamous families exhibit lower or similar fertility rates than unskilled homogamous families, in the case of heterogamous families high skilled families exhibit higher fertility than unskilled heterogamous families. In the model this come from the fact that, for similar educational incentives, homogamous families benefit from more efficient socialization technologies. Viceversa, investments in socialization in heterogamous families are bounded.

Figure 3: Fit of the model



Notes: This figure shows the average fit of the model by hj household, considering socialization rates to native culture (panel a), fertility rates (panel b), separation rates (panel c), and gains from marriage (panel d), for homogamous (red) and heterogamous families with natives (blue).

Table 5: Structural Model Parameters

	Cultural Intolerance Parameters						
c:	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah Africa	East Asia	Latin America
ΔV^c_n, Italian		24.55	60.48	76.03	85.45	82.04	24.17
$\Delta V^c_{i_E}$, Europe-EU15	16.41		52.60	4.773	88.96	90.81	14.02
$\Delta V^c_{i_O}$, Other Europe	40.32	23.05		97.14	87.65	89.96	65.01
$\Delta V^c_{i_M}$, North Africa-Middle East	71.58	7.003	98.21		62.38	60.99	83.59
$\Delta V^c_{i_A}$, Sub-Saharan Africa	59.25	96.61	83.06	43.08		58.31	54.53
$\Delta V^c_{i_S}$, East Asia	45.92	46.28	85.83	65.09	34.79		53.81
$\Delta V^c_{i_L}$, Latin America	31.65	12.18	79.85	91.32	77.51	54.09	
	Education Parameters						
Surplus for skilled children, S_c	16.67	11.05	52.95	10.69	1.38	35.52	44.61
Educational preference of unskilled parents, γ_c	1.05	1.05	1.68	1.12	0.98	1.34	0.96
	Cost Function and Extra Parameters						
Socialization Cost Parameters	σ_τ hom	16.985	Fertility Cost Parameters			σ_N hom	65.060
	λ_τ hom	0.857				λ_N hom	0.117
	σ_τ het	24.163				ϵ hom	1.028
	λ_τ het	0.451				σ_N het	82.667
Extra Marital Gain per Child	δ	0.848				λ_N het	0.195
Segregation Parameter	ρ	1.474				ϵ het	1.017
Economic complementarity in marriage	α hom	0.425					
	α het	0.427					

Notes: This table shows structural parameter estimates.

Figure 3 displays the average fit for the rate of socialization to the native culture (panel a) and the ability of the model to predict fertility rates (panel c). Overall, we match well the socialization frequencies with a correlation between predicted and observed foreign language socialization rates of 0.88, while the correlation between predicted and observed fertility rates is of 0.94.[36] In addition, the raw correlation between predicted and implied gains from marriage from the model is equal to 0.90 (panel d) and the marriage patterns observed in the data are matched very well in our empirical exercise (panel d); see Figure C.5.

5.2 Parameter estimates

Table 5 presents the estimation results.[37] Cultural intolerance parameters are estimated strictly greater than zero; that is, parents of each cultural-ethnic group have preferences for socializing children to their own cultural-ethnic group. Second, intolerance parameters are highly heterogeneous across cultural-ethnic groups, i.e., some groups are much more resilient in their cultural identity compared to others. Notably, e.g., skilled immigrants from North Africa-Middle East have very strong preferences for maintaining their cultural identity: A child integrated into the native culture, for a North African parent, is valued 70% less than

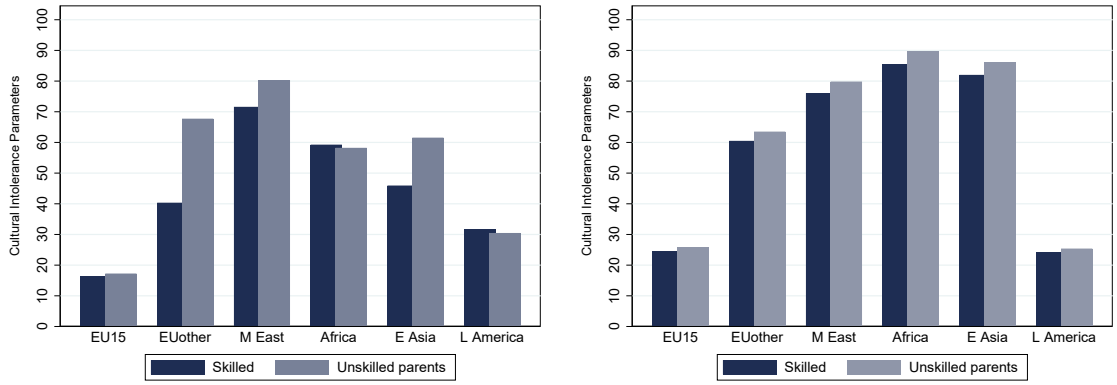
[36] Moreover, the model is also able to capture the general pattern of separation choices across groups, even though separation rates appear to be slightly underestimated.

[37] Both cultural intolerances and costs are preference parameters, hence measured in arbitrary units. But, as already noted, we normalized $V = 100$ for all cultural groups c, and hence the cultural intolerance, say of group c with respect to group c', should be interpreted as the percent reduction in lifetime utility a parent obtains if his/her child belongs to cultural-ethnic group c' rather than c. Similarly, costs are measured as percentages of the value of a child socialized to the cultural-ethnic group of the parent.

Figure 4: Cultural Intolerance Parameters

(a) Immigrants towards Natives (Demand parameters)

(b) Natives towards Immigrants (Supply parameters)



Notes: This figure reports parameter estimates for the cultural intolerance of immigrants versus natives ΔV_{is}^{ns} (panel a) and natives versus immigrants ΔV_{ns}^{is} (panel b) for all cultural-ethnic minorities i, separately for skilled (blue) and unskilled parents (grey).

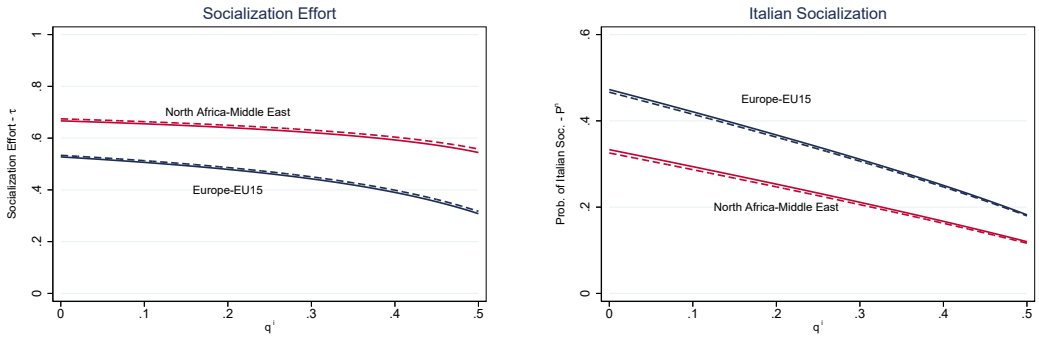
one socialized to the culture of the parent. For a skilled parent of an EU15 country, this loss is only about 17%. Cultural intolerances are similar or even higher for unskilled immigrant parents, i.e., γ_c are estimated close to one or even larger than one, especially for Other European and East Asian parents. Education parameters S_c capturing preferences for highly skilled children are also heterogeneous across groups, being higher for Other European, Latin American, and East Asian minorities in order. The cultural intolerances of Italian natives are also heterogeneous towards different minorities. To an Italian parent, a child socialized to the Sub-Saharan African cultural traits implies a 85% loss, a much larger loss than if socialized to Latin American traits, about 24%.

Third, the matrix of intolerance parameters is largely asymmetric, i.e., the intolerance of group c versus group c' is often not quantitatively close to the intolerance of group c' versus group c; see Figure 4. Notably, e.g., natives appear particularly accepting of Latin American immigrants, as we already noted; while the intolerance of Latin Americans towards natives is larger. The low estimated intolerance of natives towards Latin American immigrants is required to fit the high number of intermarriages of Italian men with Latin American women and the high fertility rate in these intermarriages (similar in magnitude to the fertility rate of homogamous Latin American marriages).

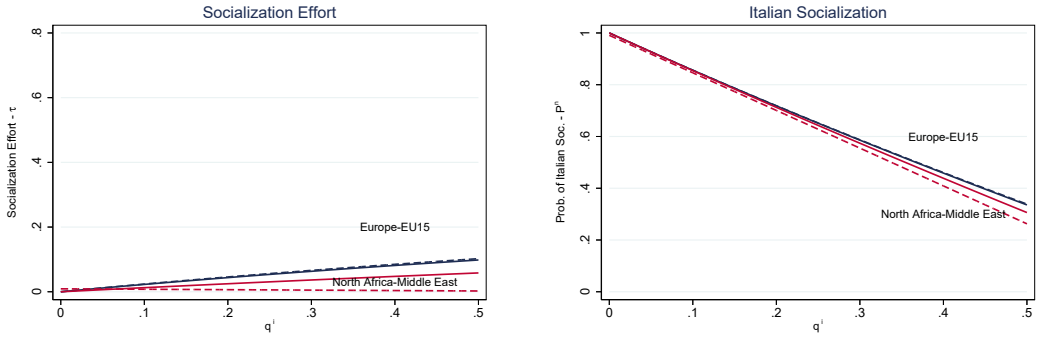
These differences in cultural intolerance preferences across groups also translate into significant differences in socialization investments. For instance, the direct socialization effort of homogamous families, τ, calculated in the extreme case in which the family belongs to a full minority i ($q^i = 0$), is estimated between 0.52 (Europe-EU15) and 0.67 (North Africa-Middle

Figure 5: Estimates of Socialization Effort and Italian Socialization Rates by Minorities

(a) Homogamous



(b) Heterogamous



Notes: This figure reports estimates of socialization effort, τ, and the Italian socialization probability, P^n, over the potential population share, q^i, for European-EU15 and North African-Middle East minorities, for homogamous (panel a) and heterogamous marriages with natives (panel b) in turn.

East), i.e., North Africa-Middle East parents have 30% higher probability of socializing children directly to their own culture, compared to Europe-EU15 parents in homogamous families. Figure 5 displays the direct socialization investment implied by our estimates, for Europe-EU15 and North Africa-Middle East (the two groups with extreme patterns), as a function of the fraction of their group in the population. For homogamous families (panel a), τ^i declines with q^i; that is, families substitute between vertical and horizontal socialization which is a consequence of the children's social interactions in the population at large. Figure 5 reports also the implied socialization probabilities. Heterogamous families with natives (panel b), instead, do not directly socialize children to the immigrant parent's culture, and the socialization effort of the native spouse τ^n increases with q^i.

Socialization costs across families are significantly different in our estimates. The cost σ_τ of heterogamous families is twice as high as the one of homogamous families.[38] This difference

[38]Costs functions are assumed to be independent of the cultural-ethnic group of spouses. The estimates

in socialization costs reinforces the gap in socialization investments between family types. We also estimate a positive difference in fertility cost σ_N, one third greater for heterogamous than for homogamous families.[39] Overall, our estimates imply that fertility investments are much more costly whenever spouses belongs to different cultural groups.

We estimate a segregation bias, ρ, of about 1.5; that is, we estimate that the contribution of society at large in the socialization process of minorities is about twice as large as the contribution implied by their actual representation in the population under random matching. This bias parameter is an indirect measure of immigrants' geographical and social segregation, e.g., in cultural-ethnic enclaves. Finally, the parameters capturing the outside option of remaining single are estimated to be highly heterogeneous both across families (homogamous vs heterogamous) and across cultural-ethnic groups, with homogamous natives showing the highest outside option parameters and the heterogamous North Africa-Middle East group showing the smallest ones.

5.3 Validation and robustness

We present in turn several different exercises to validate our estimates and we discuss some of the main assumptions of our empirical strategy, reporting on several robustness checks.

5.3.1 Validation

The first validation exercise is focused on our main parameters of interest. We externally validate our cultural intolerance estimates (a revealed measure of perceived cultural distance) across cultural-ethnic groups by relating them to various cultural distance measures commonly used in the literature (Spolaore and Wacziarg, 2009, 2016). Figure 6 documents a systematic positive correlation between our cultural intolerance estimates and various cultural distance measures. These cultural measures explain from 15 to 30 percent of the variation in cultural intolerance. We find it remarkable as cultural distance measures are, by construction, symmetric, while our estimates do not impose any restriction on symmetry. Weighted estimates by the number of marriages per match report a systematically higher correlation, from 40 to 70 percent.

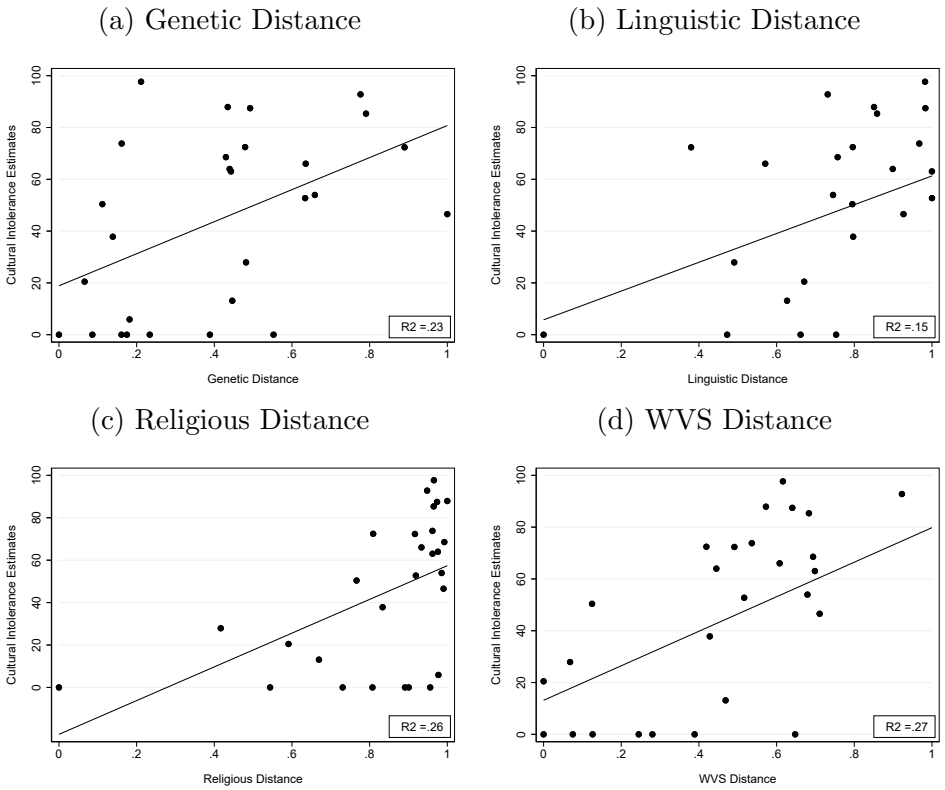
Our second validation exercise consists in exploiting our preference parameters to predict the distribution of marriages observed for (out-of-sample) newly formed marriages from 2013 to 2019.[40] The model fits out-of-sample marriage data very well; Figure 7 shows the

of λ_τ, associated to the degree of convexity of costs, are comparable across family types.

[39] The parameter ξ is also higher for heterogamous families (about 20%).

[40] Because of data availability, in this exercise, we focus on the distribution of marriages along cultural-ethnic lines, and we disregard the educational variability. Theoretically, we impose $\gamma_c = 1$ and $S_c = 0$, and

Figure 6: Cultural Intolerance Estimates and Cultural Distance Measures



Notes: The figure shows the relationship between our cultural intolerance estimates and various measures of cultural distance: cultural distance along genetics (panel a), language (panel b), religious (panel c), and values (panel d). Data are available thanks to Spolaore and Wacziarg (2016).

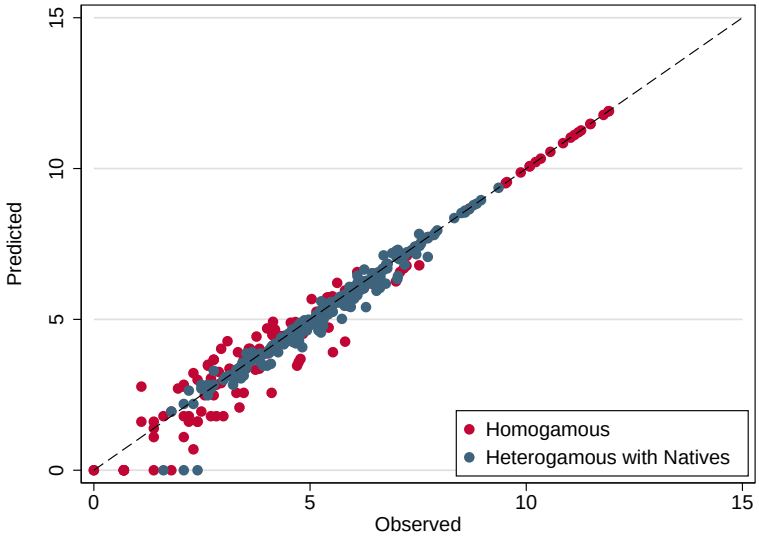
relationship between the number of marriages observed from 2013 to 2019 and the number of marriages predicted by our model by match and region. Figure C.6 displays the distribution of men and women along cultural lines in the sample and out-of-sample period.

The third validation regards the socialization rates of divorced couples, which we compare with socialization rates observed in our survey data but not targeted in the estimation. The model matches these external moments very well, the correlation between the observed and predicted rates of socialization to the native culture is equal to 0.76; see Figure C.7.

In the last exercise, we validate our estimates by asking whether they predict well the distribution of marriages by province, a higher level of geographical disaggregation; see Figure C.8. We exploit variation in the distribution of cultural-ethnic traits across provinces within the same region. Indeed, we observe variation in the distribution of cultural-ethnic traits across provinces within the same region (within region heterogeneity accounts for about 1/3 of the total observed variation) and we show that our estimates, based on regional moments,

re-estimate the other parameters. We exploit these same estimates also for the other two validation exercises.

Figure 7: Model Validation - Non Targeted 2013-2019 Marriages by Match and Region



Notes: This figure displays the scatterplot of the relationship between the number of marriages observed in the out-of-sample data (in log) in the years from 2013 to 2019 and the number of marriages predicted by the model (in log) by region for homogamous families (red) and heterogamous families with natives (blue). Out-of-sample aggregate marriage data doesn't provide details on heterogamous marriages between different immigrant groups.

in fact predict successfully the pattern of marriages at this more granular level.

5.3.2 Robustness

We discuss our main assumptions in turn.

Cultural-ethnic socialization is proxied with language socialization. Ethnic identity and spoken language are relevant culturally related specific attributes and both allow the direct transmission of cultural characteristics across generations.[41] We measure socialization by the *language spoken at home* within the family, as a form of parental cultural investment. While, reasonably, all children living in Italy learn Italian at school, speaking Italian at home when at least one spouse is an immigrant, in our interpretation, reveals deep-seated preferences for integration relative to ethnic identity (Bazzi et al., 2019; Salari, 2020). To corroborate this interpretation, we provide evidence along two directions. First, we document that our measure of Italian linguistic socialization influences the achievement and educational choices of immigrant students. We obtained student-level data on standardized test scores in reading and math administered by INVALSI to all students in Italy at the end of grade 5. The test

[41] See Casey and Dustmann (2008); Ginsburgh and Weber (2011); Clots-Figueras and Masella (2013); Fouka (2020). Schwartz (2013), in particular, underlines the parallel between ethnic and linguistic homogamy.

is identical for all students in a given grade, and it is blindly scored, hence results are fully comparable across schools. Crucially for us, INVALSI data also contain survey information on the main language spoken at home by students, as well as rich baseline information on students and family background. We describe in details the Italian educational setting and INVALSI data in Appendix B. Table 6 shows that speaking Italian at home improves the achievement in test scores of immigrant students both in reading and math. Immigrant students who speak Italian at home exhibit higher achievement by a 0.20 (0.11) standard deviations in reading (math). The point estimates are significant and relevant in magnitude. Moreover, by exploiting the longitudinal structure of the data, we investigate the language long-term impact on students' educational careers. We show that speaking Italian at home drives students into high demanding schools, i.e., Italian socialization at home (during the elementary school period) increases the probability of attending an academic or technical high school (as opposed to vocational one) by 2.7 percentage points, on a baseline rate of 82 percent for immigrant students in grade 10. This might ultimately have long-term implications for access to college and occupational careers (Brunello and Checchi, 2007; Carlana et al., 2021). Results are robust to various checks, see Appendix Table C.10.

Second, we provide survey evidence that Italian use at home is associated with weaker ethnic identity and stronger social integration networks, educational achievement and aspirations, and proficiency in the Italian language. Appendix Table C.11 shows that children who speak Italian at home are: (i) about 50% more likely to have Italian friends (outside school), and are more likely to speak Italian with school mates and friends (columns 1-3); (ii) 15% more likely to achieve high education, to pass the academic year and have greater aspirations for their future educational career path (columns 4-6); (iii) significantly more

Table 6: Italian Language Socialization and Educational Outcomes

	(1)	(2)	(3)
Dep var.:	Reading std test score, 5th grade	Math std test score, 5th grade	High-track choice, 10th grade
Italian at Home	0.203***	0.109***	0.027***
	(0.004)	(0.004)	(0.003)
Observations	330,739	345,980	90,656
R-squared	0.144	0.098	0.058
Province & Cohort FE	Yes	Yes	Yes
Student Controls	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes

Notes: This table shows how our measure of Italian linguistic socialization (Italian at home) influences the achievement and educational choices of immigrant students. The dependent variables include the reading or math standardized test score for students in grade 5 (columns 1 and 2), respectively, and a dummy equal to one for students attending high-track (academic or technical) schools in grade 10 and zero otherwise (column 3). Test scores are standardized with zero mean and standard deviation of one. The sample includes all students with at least one immigrant parent. All regressions include student controls (gender, regular schooling, a dummy for first generation immigrants, and a dummy for kindergarten), family controls (mother's and father's education and a set of dummies for socio-economic background), as well as province and cohort fixed effects. Robust standard errors clustered at school level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

proficient in Italian in writing, reading, speaking, and comprehension (columns 7-11).[42]

The geographical unit of reference for marital and intra-household choices coincides with the region. This might appear problematic if a higher level of disaggregation would reveal different patterns of segregation of minorities across geographical units. In fact, the existence of ethnic enclaves within regions is consistent with our empirical strategy if residential segregation represents in itself a costly mechanism of cultural socialization, in line with Bisin and Verdier (2000) and Bisin et al. (2004). The validation exercise in the previous section, exploiting the higher provincial level of geographical disaggregation, confirms this point.

Differences in fertility might be due to systematic variation in observables across households. To partial out these differences, we regress fertility on a set of covariates including marital duration, age at marriage of spouses, educational attainment, employment and occupational standing.[43] We then compute the fertility moments using the residual variability out of this linear regression model, thus unrelatedly to potential confounding in our estimation. Figure C.9 reports cultural intolerances in line with our main estimates, while only the demand of cultural identity of East Asian minorities appears underestimated at baseline.

The distribution of the population across regions is exogenous. We abstain from modeling endogenous moving and/or residential location decisions. Endogenous moving or location decisions would be problematic for our estimates, if these decisions were motivated in part by marriage and socialization, as well as by unobserved heterogeneity. Consider the natural hypothesis that minorities that are particularly attached to their cultural identity choose to locate into more segregated areas. In this case, we would expect a positive correlation between vertical and horizontal socialization. Figure C.10 describes this relationship by plotting the probability of direct socialization in homogamous families over the corresponding population share by regions, showing instead a negative and, at times, statistically significant relationship. The evidence is consistent with direct cultural transmission within the family substituting horizontal socialization, i.e., if anything, minorities in more segregated regions display lower socialization rates to the language of parents. This substitution pattern, thus, represents a lower bound, net of complementarity in residential selection. To further alleviate potential concerns of endogenous location choices, we rely on pre-determined settlements of ethnic groups of immigrants across regions. Specifically, we predict the population distribution by ethnic group and region, by exploiting pre-existing variation in the geographical

[42]Relatedly, other studies uncover a positive association between the proficiency in the destination language and socio-economic integration, e.g., favoring the educational achievement of lag-behind children (Dustmann et al., 2010), as well as employment and earning opportunities (Dustmann and Fabbri, 2003).

[43]Our focus is primarily on fertility, as differences in separation rates are less salient for our identification. For further evidence on separations, see Tura (2020) who documents that intermarriages exhibit a 16% higher risk of separation compared to homogeneous marriages.

distribution of immigrants by ethnic group observed in 1993, interacted with subsequent inflows by origin, as in a shift-share instrument strategy (Card, 2001; Tabellini, 2020).[44] Indeed, the distribution of immigrants exhibit a strong geographical clustering along ethnic lines, and settlement patterns of ethnic minorities are a good predictor of location choices of newly arriving immigrants. Results are in line with our baseline estimates.

6 Long-run integration patterns

In this section, we simulate the dynamics of the distribution of cultural-ethnic traits in the population induced by our structural model of marital matching, fertility, divorce, and socialization. While the exercise rests on the strong assumption that parameters are invariant over time, these simulations should be interpreted to highlight the implications of our estimated model with respect to the prospective pattern of cultural-ethnic integration of different minorities in Italy.[45] Moreover, the notion of integration we adopt, given our data, refers to the practice of speaking Italian at home; that is, an individual belonging to a minority is integrated in our simulations when living in a household speaking Italian at home.

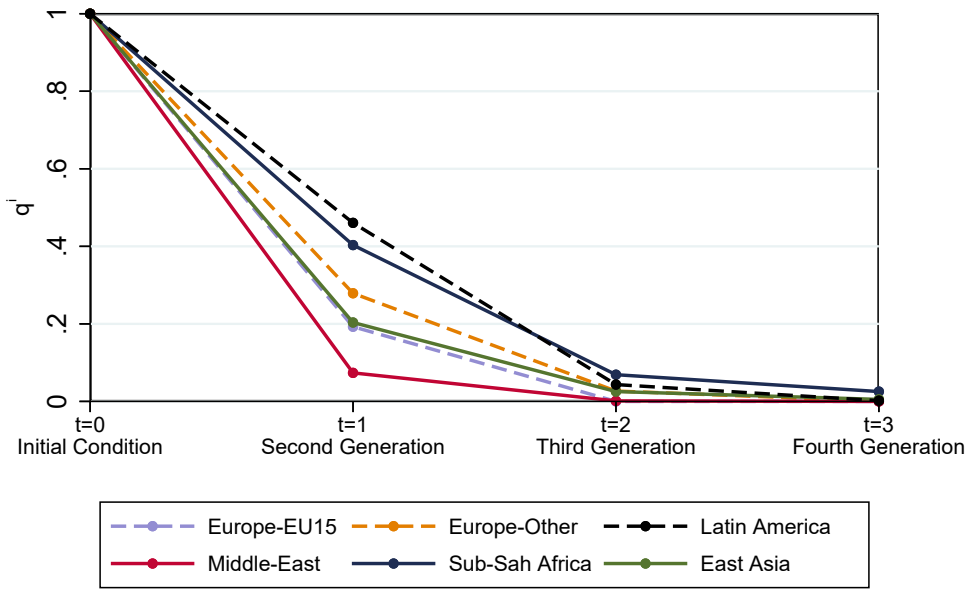
6.1 Population dynamics

The time unit in the simulations is a generation, i.e., a time interval of about 25-30 years. We fix the initial condition, generation $t = 0$, to coincide with the distribution of the population by region and ethnic group in our data. More precisely, while we observe the demographic characteristics of the Italian population over time, from 1995 to 2012, we interpret them for these simulation exercises as representing a cross section of the population in 2012, by region and ethnic group. Let this distribution be denoted p_t. The structural model we have estimated induces a map from p_t into p_{t+1}. Indeed, the model maps any distribution p_t into a vector of demographic characteristics of the population at time t, in terms of marital matching, fertility, divorce, and socialization, by ethnic group and region. The mapped fertility and socialization at t, by region and ethnic group, induces, in turn, a distribution

[44] Notice that, as anticipated in the Introduction, immigration is a relatively recent phenomenon in Italy, and the number of (legal) foreigners residing in Italy in 1993 was about 630,000.

[45] The model might be generalized to allow for inter-temporal dependence of cultural intolerances, relaxing the assumption that cultural intolerance parameters are constant across generations. For instance, the cultural intolerance of second generations could be set equal to $\eta \Delta V_{is}^{ns}$, where $\eta \in [0, 1]$ is a factor capturing the distance of second generations with respect to the minority culture, due for instance to a lower of natives stereotypes towards minorities or faster integration. However, empirically, we cannot identify η from the single cross-section of marriage data we have. Identification would need to rely on a repeated cross-section of marriage data for first and second generations.

Figure 8: Long-run Dynamics of Cultural Traits (index=1 in $t = 0$)



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations.

of the population of the children of the population at time t, p_{t+1}.[46] The same simulation procedure, recursively, induces $p_{t+2}, p_{t+3}, ...$[47]

The simulated long-run dynamics of the fraction of the population with cultural-ethnic trait i for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$ are reported in Figure 8, normalized so that $q^i_t = 1$ in $t = 0$ for comparability.[48] See Figure C.11 for non-normalized long-run dynamics of cultural traits. Despite estimates of cultural intolerance highlight immigrants' strong preferences for maintaining their cultural identity, all cultural-ethnic minorities are simulated to integrate into the Italian majority along the language dimension. Indeed, the *integration rate*, defined as the reduction in the fraction of the total population (immigrants and natives) which is composed of immigrants who are not integrated to the native Italian culture, is 75% over

[46]Reproduction is asexual in the model, hence we consider future generations populated by men and women of equal proportion. Note also that the individuals in the population composing the distribution p_t are distributed across the age dimension. We disregard this in the estimates, and hence also in the simulation, but we can interpret the distribution p_{t+1} as representing the same distribution across age.

[47]At each step, we compute the equilibrium in the marriage market subject to feasibility constraints, represented by Equation (2). This amounts to solving a system of $2 \times 2K$ quadratic equations in as many unknowns, with Z the number of cultural traits in the population and two education categories, for each of the R regions. To this end, we take advantage of an iterative projection fitting procedure (IPFP) designed to find projections on intersecting sets of constraints, by projecting iteratively on each constraint (Galichon and Salanié, 2021; Galichon, 2018).

[48]The distribution of cultural traits in the population at time $t = 0$ is computed from population data as the average across regions weighted by the total resident population. The cultural belonging of first-generation immigrants is identified by the country of origin.

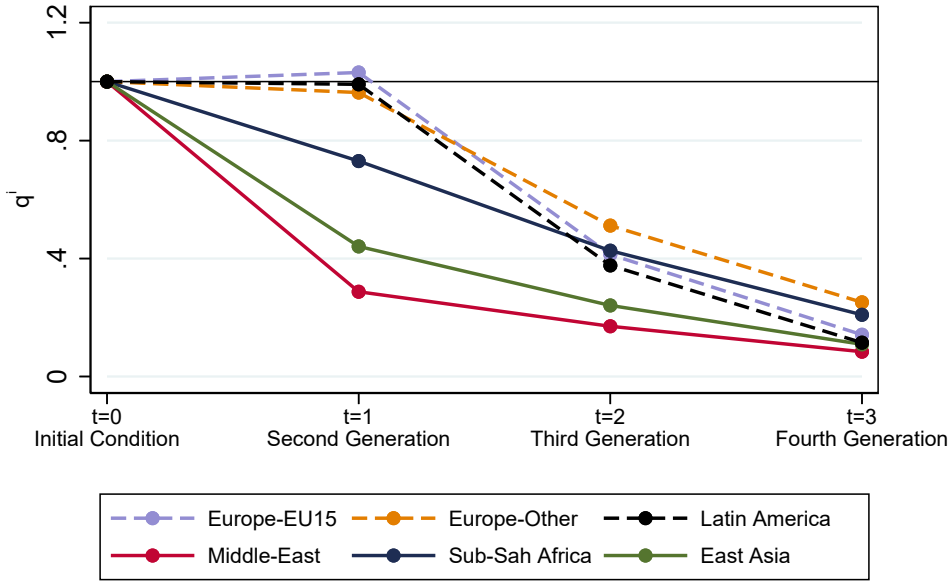
the period of a single generation.[49] The rate of integration is however heterogeneous across cultural-ethnic groups. In particular, the North African-Middle Eastern minority integrate almost completely, 92%, in a single generation. A similar patterns of convergence is also displayed by the European-EU15, East Asian, 80% in one generation, and Other European minorities, 72%. On the other hand, a significantly slower integration rate is achieved by the Latin American minority, whose generation at $t = 1$ exhibit an integration rate of 54% and reaches full convergence only by the fourth generation. A slower integration rate also characterizes the and Sub-Saharan African minorities, 60% in one generation. For third generations overall 97% of immigrants converge towards Italian culture in about 50-60 years. If we disregard education, assume $S_h = 0$, the overall integration rate of immigrants remains unaffected and equal to 75% over the period of a single generation. However, the heterogeneity in integration rates across cultural-ethnic groups is affected. For instance, East Asian minority exhibits a slower pattern of convergence induced by higher estimated fertility.

The patterns of cultural integration of European-EU15 and Other European minorities are the result of their relatively low cultural intolerance preferences. In a similar way, the East Asian and Sub-Saharan African minorities' slower integration is due in part to their higher intolerance parameters. But intolerance parameters are not the only determinants of the dynamics of integration of different cultural-ethnic groups. Homogamous marriage rates, fertility rates, and other demographic characteristics in fact turn out to have sizable independent effects on cultural integration in the simulations.[50] This is clearly illustrated by the fact that, while North Africa-Middle East, Sub-Saharan Africa, and East Asia show relatively comparable cultural intolerance preferences, they display significant differences in the dynamics of integration. Indeed, a strong estimated selection into homogamous marriages of immigrant from Sub-Saharan Africa allows them to sustain their cultural heterogeneity by accessing superior direct socialization technologies. On the other hand, fertility rates are particularly high for East Asian minorities and this is a fundamental factor behind this minority's integration pattern. Finally, the relative success of Latin America in securing their cultural distinctiveness over time is due in large part to the fact that they turn out to be uniquely able to socialize children also in heterogamous marriages with natives.

[49] The integration rate of a specific cultural-ethnic group is then the reduction in the group fraction in the total population, over successive generations. This notion of integration rate differs from the rate of socialization to the native culture, i.e., the fraction of second-generation immigrants (born from marriages with at least an immigrant spouse) speaking Italian at home, which is 85% overall in the simulations.

[50] With respect to fertility, this is the case even though predicted fertility rates for all groups are below reproduction level, which potentially has implications on marriage market competition as well.

Figure 9: Long-run Dynamics of Cultural Traits with Italians Fully Tolerant



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups (index=1 in $t = 0$), over successive generations assuming the case of complete tolerance of Italian majority towards minorities, $\Delta V_n^i = 0$.

6.2 Counterfactual cultural intolerance parameters

In this section, we examine more in detail the mechanisms that promote the cultural integration of immigrants. In particular, we analyze the role of cultural intolerance parameters, studying the dynamics of the distribution of cultural-ethnic traits in the population under several counterfactual values of ΔV_h^k, for $h, k \in C$. We connect the results arising from our counterfactuals with reduced form evidence in the literature.

We consider, first, the case in which $\Delta V_{ns}^{is} = 0$ for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$; that is, we consider a counterfactual environment in which natives are fully tolerant towards minorities, and offer complete acceptance of the immigrants' cultural diversity. Results, displayed in Figure 9, show that this counterfactual experiment induces on average a 25% reduction in the fraction of the population composed of non-integrated immigrants after one generation. In the long-run, immigrants start integrating but slowly, so that the heterogeneity in the cultural traits in the population shows a remarkable persistence. While immigrants maintain strong preferences for cultural identity, an increase in the acceptance of natives has the effect of making intermarriages of immigrants with natives more valuable for both. Indeed, the coordination failure in children socialization that in the general framework undermine the value of intermarriages is now muted, and the immigrant spouse is able to socialize children to his/her own trait. In our counterfactual simulation, we observe a large increase

in intermarriages with natives, and in parallel a lower demand for homogamous marriages. Furthermore, fertility rates in intermarriages with natives are substantially increased. Socialization to the Italian language is reduced, as it is driven only by horizontal socialization of society at large. All these effects induce a reduction in immigrants' integration compared to the baseline; see Figure C.13 and Figure C.14 for a representation of the mechanisms driving the dynamics of integration. The dynamics of integration is still heterogeneous however, across groups. Specifically, we find that, by the fourth generation, the integration rate of Other European, Sub-Saharan African, and North Africa-Middle East is about 74%, 78%, and 90%, respectively.

We study also the extreme opposite case in which Italians are fully intolerant towards all minorities, $\Delta V_{ns}^{is} = 100$ for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$. The dynamics of integration of ethnic minorities follow the results in the baseline, with an integration rate of about 80% overall over a single generation; see Figure C.15.

These results are in line with those obtained by Fouka (2020), who examines the immigrants' response to the shift in attitudes towards immigration in the US in the aftermath of WWI, which lead to imposition of language restrictions in elementary schools. She shows that individuals subject to language bans i) were less likely to volunteer in WWII, ii) were more likely to marry within their ethnic group, and iii) strengthen the vertical socialization of their children (high foreign names).[51] In our model, this shift in attitudes maps into an increase in ΔV_n^i (lower supply of acceptance on the part of natives), while the language ban maps into higher costs of socialization. According to our counterfactual analysis, this would induce higher homogamy in marriages and higher socialization rates to the culture of minorities, in line with Fouka (2020)'s results.

To study the potential role of economic incentives for integration, we consider the case in which ΔV_{is}^{ns} increases for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$ by 20% of ΔV_{ns}^{is}, which could reveal either a reduction in economic incentives to integration (e.g., the job opportunities available for the more integrated immigrants) or stronger cultural intolerances. When we simulate our model in this case, the dynamics of integration of minorities towards Italian culture accelerate by 10 percentage points compared to baseline estimates. Results, displayed in Figure 10, appear counterintuitive. In principle, we might have expected that, in the absence of economic incentives, the strengthening of migrants' cultural preferences would make their cultural integration more demanding and hence slower, by providing higher socialization rates and reducing heterogamous marriages. On the contrary, we show that the reduction

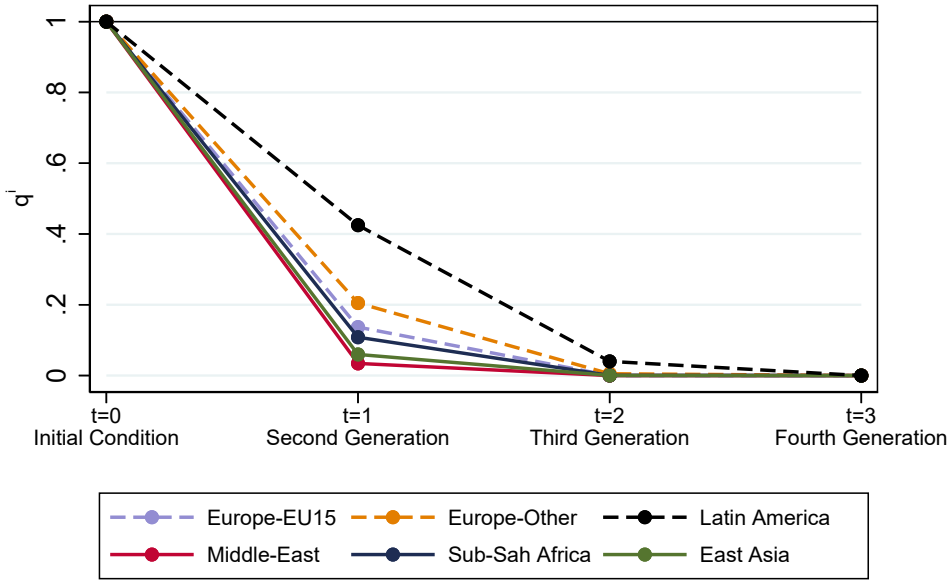
[51] This documented backlash in terms of parental investment into their children's identity does not necessarily imply a clear-cut sign in the overall effect on socialization outcomes as two opposite forces are at play: higher socialization to native culture via schooling, counteracted by higher socialization of the family to the minority culture.

in economic incentives to integration contributes to accelerate cultural convergence.

Indeed, the stronger attachment of minorities to their identities makes marriage riskier and costlier, by commanding a larger investment of resources in socialization. The value of marriage significantly reduces for both homogamous marriages of immigrants and heterogamous marriages (while, the value of homogamous marriages of natives remains unchanged), leading to a general equilibrium effect of compression of the marriage market for immigrants. Hence, cultural convergence in this setting is induced by a lower participation in the marriage market and lower fertility of immigrants. In other words, in this counterfactual, the acceleration of integration is mostly an effect of the reduction in the population growth for immigrants with respect to natives. The probability that a child with an immigrant parent is integrated to the Italian culture is lower in the absence of economic incentives (i.e., the parental socialization effort becomes more effective), but the fraction of the total population (immigrants and natives) belonging to the native culture, on the contrary, is higher.[52]

These results are in accordance with those of Adda et al. (2020), who study how marriage and separation choices of immigrants respond to a change in the labor market value of marriage. Specifically, Adda et al. (2020) study the effects of the EU enlargements in

Figure 10: Long-run Dynamics of Cultural Traits net of Economic Integration Incentives



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups (index=1 in $t = 0$), given a proportional increase in ΔV_i^n for all minorities by 20% of ΔV_n^i.

[52]Other counterfactuals changing demand cultural intolerances are discussed in the Appendix Figure C.16.

2004 and 2007 which,[53] by guaranteeing permanent legal status and the right to work to the citizens of the new Eastern European member-states, eliminate their prior benefits of marrying an EU native. This paper shows that, after these enlargements i) the probability of the relevant intermarriages decreases, and ii) the hazard rate of separation of these intermarriages increases. In our model, this experimental setting maps into a reduction in economic incentives to integration, a larger ΔV_{is}^{ns}, which translates into the effects on intermarriages and separations observed by Adda et al. (2020) after the enlargements.

7 Policy and welfare

Cultural integration of immigrants in our structural model of marriage, fertility, and socialization is the equilibrium outcome of two interacting mechanisms, frictionless matching and collective decisions at the household level. Both these mechanisms are efficient.[54] A *dynamic* welfare analysis of equilibrium in our model however displays an externality in the fertility and socialization choices of households. Individual households at time t in fact do not take into account the aggregate effects of their choices on the size and the composition of the population by cultural-ethnic group in the future, from time $t+1$ onwards. A dynamic welfare analysis of policy interventions in our model can be performed with a utilitarian social welfare function W, which aggregates the utilities of all (groups of) agents in the economy, in different family types and ethnic groups, over multiple periods. We report next on such welfare analysis for a specific policy choice, the social segregation of immigrant minorities.

Consider the parameter ρ in our model, which we introduced to represent the strength of the immigrants' cultural-ethnic network through a segregation bias in their pool of reference. We consider ρ as, at least in part, a policy variable. In fact, geographic and social segregation of immigrants (e.g., in living quarters, schools, religious gathering, social clubs, etc.) is the result of their own choices as well as of governmental and local policies regarding e.g., public housing, freedom of religion, schooling, local police, etc. We have estimated ρ of about 1.5 in our empirical analysis. We simulate now the effects of policies which change ρ to a range of values, from 1 to 15, at time t, before agents match in the marriage market. While we solve for the whole equilibrium at t and $t+1$, we concentrate here on the welfare effects of the segregation bias ρ. At time t, as ρ increases, immigrants benefit from a more efficient socialization technology (via the horizontal channel), and hence homogamous marriages increase largely in value (i.e., expected marital utility) and in number. Instead

[53] For further details, we refer to footnote 8.

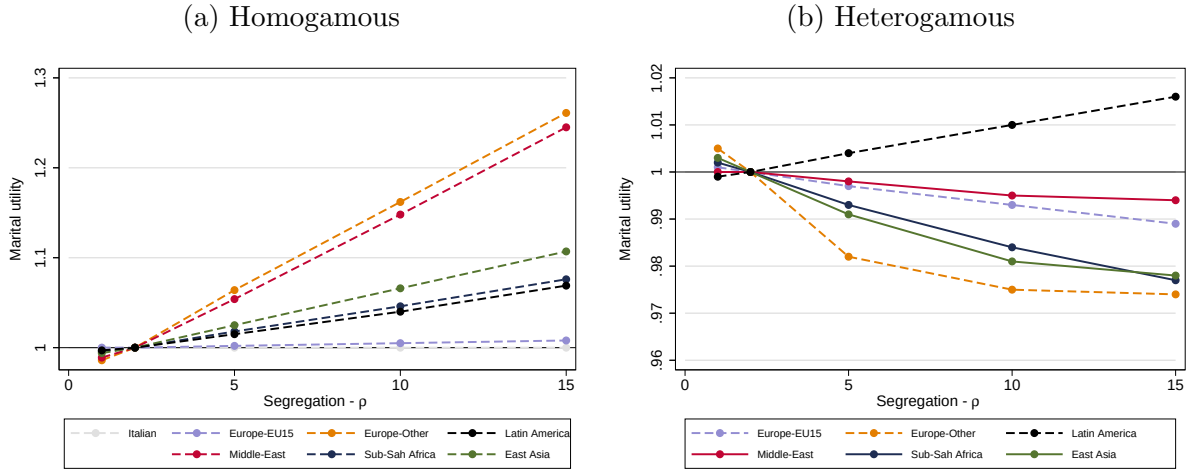
[54] To be precise, the collective decision problem has a possibly inefficient component in that socialization after separation in divorce is chosen non-cooperatively by parents.

by exacerbating the socialization conflict, social segregation decreases the value and hence the number of heterogamous marriages with natives.[55] Figure 11 reports the marital utility response to an increase in ρ for homogamous (panel a) and heterogamous families (panel b). Finally, homogamous marriages of natives also increase in number via market equilibrium, even though they are not affected in value as they socialize their children at no cost; that is, independently of ρ.

Panel a. of Table 7 reports the number of marriages and the utilitarian welfare computed at time t, by family type (columns 1-6). Social welfare increases monotonically with segregation bias ρ for homogamous marriages of both immigrants and natives, while the heterogamous marriages display a welfare loss. We also report welfare results at the individuals level distinguishing natives from immigrants, i.e., lumping immigrants of all groups into a single average category (columns 7-8 of Table 7). Under natural assumptions,[56] segregation ρ proves to be welfare improving for any weight structure, as both natives and immigrants realize welfare gains.

The most interesting effects appear at time $t+1$, when the externality in the fertility and socialization choices of households plays a fundamental role. As represented in Figure 12,

Figure 11: Marital Utility Response to Segregation at t, by Family Type



Notes: This figure reports the average marital utility response rate to an increase in segregation bias ρ, computed at time t, by ethnic-group for homogamous (panel a) and heterogamous marriages with natives (panel b). Marital utility rate index to 1 at baseline ρ.

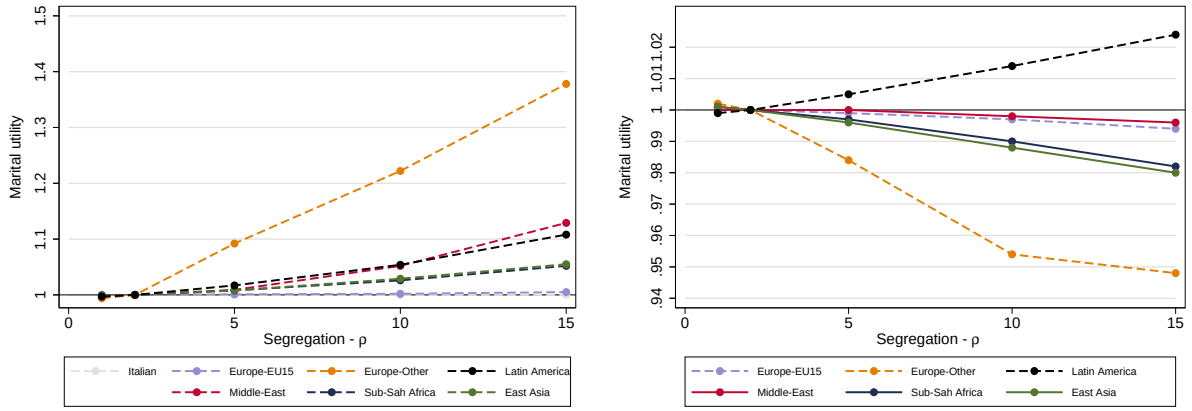
[55] Heterogamous marriages of Latin Americans with natives represent an exception, their marital utility increases with the segregation ρ, as Latin Americans are the only minority who is able to socialize children also in heterogamous marriages with natives.

[56] Under transferable utility, the aggregate value of social welfare cannot be identified without additional data on transfers between spouses in heterogamous marriages. But it is natural to assume that utility in these marriages is transferred proportionally to the spouses outside options, which we have estimated in our empirical analysis (see also Adda et al., 2020).

Figure 12: Marital Utility Response to Segregation at $t+1$, by Family Type

(a) Homogamous

(b) Heterogamous



Notes: This figure reports the average marital utility response rate to an increase in segregation bias ρ, computed at time $t+1$, by ethnic-group for homogamous (panel a) and heterogamous marriages with natives (panel b). Marital utility rate index to 1 at baseline ρ.

the marital utility of homogamous (panel a) and heterogamous families (panel b) calculated at time $t+1$ respond to segregation in a similar direction as at time t. But, the fertility

Table 7: Social Welfare Response to Segregation by Family Type and Ethnic Group

	Family Types:						Ethnic Groups:	
	Immigrant Families		Native-Migrant Families		Native Families		Natives	Migrants
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Segregation bias	Marriages	Welfare	Marriages	Welfare	Marriages	Welfare	Welfare	Welfare
	a. Aggregate Social Welfare W_t at time t							
ρ =1	63313	216	507272	1539	3578365	30606	31491	870
ρ =1.5 (baseline)	71347	247	489352	1485	3582663	30643	31494	881
ρ =5	120798	458	413888	1267	3600125	30792	31506	1012
ρ =10	189460	790	328210	1027	3619386	30957	31520	1253
ρ =15	225817	985	279174	890	3630407	31051	31531	1396
	b. Aggregate Social Welfare W_{t+1} at time $t+1$							
ρ =1	1945	7	26918	76	890658	7618	7664	36
ρ =1.5 (baseline)	2684	9	31096	87	888939	7603	7657	43
ρ =5	6313	24	42736	119	884220	7563	7637	69
ρ =10	27118	125	38923	108	883829	7560	7626	166
ρ =15	46883	243	31387	90	883738	7559	7611	278

Notes: This table shows the number of marriages and the aggregate value of social welfare (in 10,000) by changing the segregation parameter ρ to a range of values, from 1 to 15. Panel a. reports marriages and the value of social welfare W_t computed at time t, while panel b. reports social welfare W_{t+1} at time $t+1$. In columns 1-6, the welfare function W aggregates the utilities of agents by family type into immigrant families, native-migrant families and native families. Columns 7-8 report the welfare function W, distinguishing natives and immigrants.

and the socialization rate in homogamous immigrants households at time t grows steeply with ρ affecting the size and composition by cultural-ethnic group of the population at $t+1$. As ρ increases then, the larger fraction of second-generation immigrants available in the marriage market increases the number of both homogamous immigrant and heterogamous marriages, in turn increasing their contribution to the social welfare function. Panel b. of Table 7 reports the marriage and welfare effects computed at time $t+1$. Maintaining the assumption that utility is transferred proportionally to the spouses outside options, we have that social welfare increases with segregation bias ρ for any weight structure, as overall both natives and immigrants realize welfare gains. The increase in heterogamous marriages compensate in the aggregate the reduction in the contribution of homogamous marriages of natives (which have constant value but decrease in numbers) to the social welfare function.[57] In consideration of the demand of minorities to preserve their cultural identity, segregation policies lead to positive short-run and dynamic welfare effects.

8 Counterfactual migration inflows

In the last few years, Italy has experienced a significant increase of migration inflows, mainly originating from Sub-Saharan Africa and Middle-East countries. We study the effects of such a rise in immigration on cultural heterogeneity in Italy, by performing two counterfactual exercises. In both cases, we exogenously increase the number of second-generation minorities and study the long-term implications of this increase on cultural convergence. In the first exercise, we double the share of second-generation minorities *proportionally for all cultural-ethnic groups*; while in the second exercise we still double the overall share of second-generation immigrants by *assigning one third of the increase exclusively to North African-Middle East, Sub-Saharan African and East Asian minorities.*

Keeping constant the share of each group, Figure 13 compares the dynamics of the distribution of cultural-ethnic traits in the population at the baseline (solid line) with the distribution resulting from the rise in immigration (dashed line). Overall, doubling the shares of second-generation minorities at $t = 1$, leads to a reduction in the integration rate of 10 percentage points for third generations, 87% compared to 97% at the baseline. More in detail, the rise in migration inflows has no effect on the cultural integration of European-EU15 and North African-Middle East minorities already in the third generation. On the contrary, the incoming waves of Other European, Sub-Saharan Africa and East Asia immigrants produce an effect in delaying their full convergence to host country cultural practices. In particular, we estimate a 11 and 5 percentage points reduction in the integration rate of Sub-Saharan

[57] The marriage market endogenously increases with ρ at time $t+1$, while the number of singles reduces.

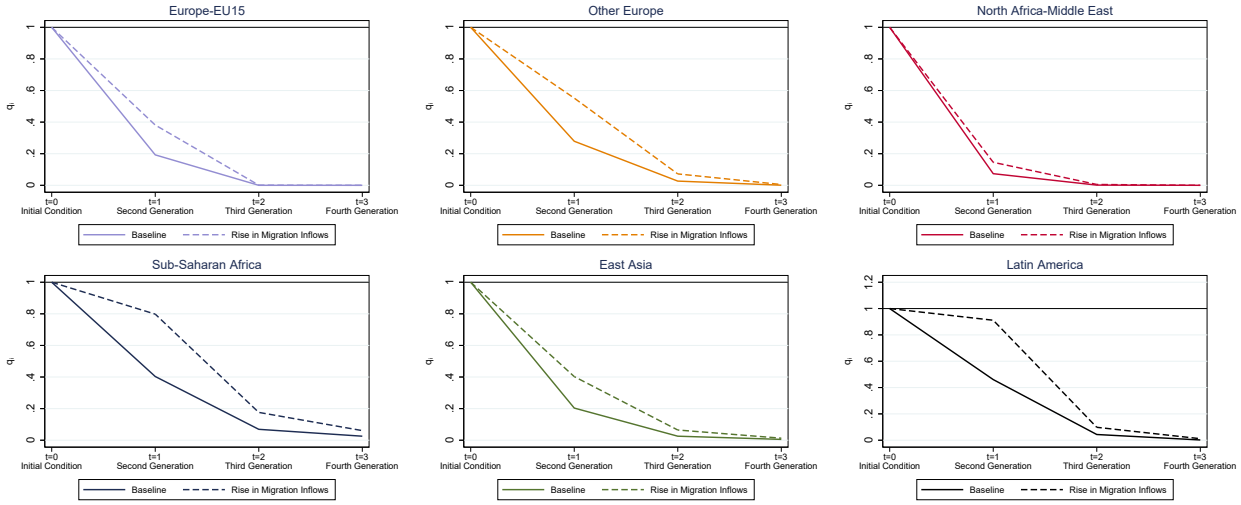
African and East Asian minorities, respectively, compared to baseline.

We observe qualitatively similar results (even though stronger in magnitude) in the second exercise, when we modify the relative distribution of second generations, overweighting North Africa-Middle East, Sub-Saharan Africa and East Asia; Figure 14 displays the resulting integration response. For a comparable increase in migration flows, the three groups highlight significant differences in integration patterns, with Sub-Saharan African and East Asian minorities accentuating their successful transmission of cultural values dramatically. In particular, the cultural integration response of North Africa-Middle East immigrants to the exogenous rise in inflows is reduced by only a 6 percentage points. The response of East Asian and Sub-Saharan African minorities to a comparable variation ranges from 10 to about 30 points, slowing down the process of cultural integration.

9 Conclusions

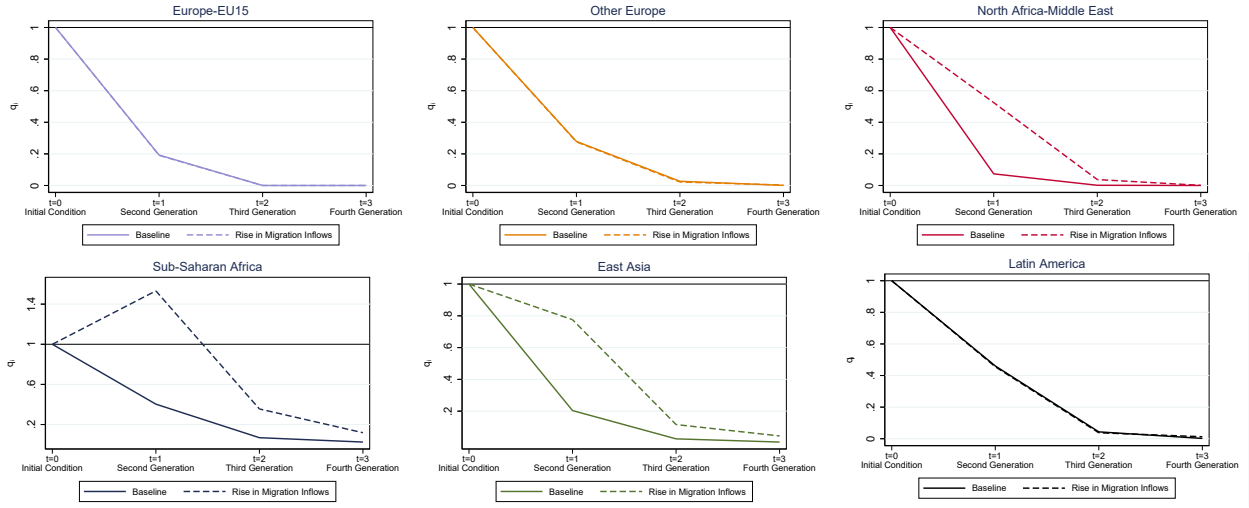
As cultural boundaries are increasingly salient, the design of adequate and successful policies to integrate minorities is a fundamental and challenging policy objective of modern societies. In this paper, we offered a new perspective to interpret cultural integration as an equilibrium outcome of marital matching and collective household decisions. We show by counterfactual analysis how the dynamics of immigrants integration over time respond to variations both

Figure 13: Long-run Dynamics with Proportional Increase in Migration Inflows



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations, normalized so that $q_t^i = 1$ in $t = 0$. The solid line represents the dynamics at baseline, while the dash line represents the dynamics after doubling the share of second generation immigrants, proportionally for all minority groups. Figure C.17 reports the non-normalized long-run dynamics.

Figure 14: Long-run Dynamics with Compositional Increase in Migration Inflows



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations, normalized so that $q_t^i = 1$ in $t = 0$. The solid line represents the dynamics at baseline, while the dash line represents the dynamics after raising the share of North Africa-Middle East, Sub-Saharan Africa and East Asia second generation immigrants. Figure C.18 reports the non-normalized long-run dynamics.

in the demand of immigrants for the preservation of their cultural identity, as well as in the supply of acceptance of the immigrants' cultural diversity on the part of natives. These findings have in principle novel implications for the evaluation of different immigration policies, beyond across-the-board integration on one side and restrictive closed-border policies, on the other. Indeed, several results we obtain might not clearly emerge without an integrated equilibrium analysis of matching, fertility and socialization: i) an higher acceptance of the culture of minorities on the part of natives reduces the integration rate of immigrants; while ii) a reduction in economic incentives to immigrants instead accelerates their integration rate.

Moreover, by examining how family investments and the social environment interact to shape the cultural identity of new generations, we show that a policy that strengthens the immigrants' cultural-ethnic network through a segregation bias in their socialization pool enhances social welfare. This conclusion might represent a starting point for a debate about e.g., residential location, school choice, and religious freedom of immigrants, with far-reaching implications, as societies become more ethnically heterogeneous.

References

Abdelgadir, A. and V. Fouka (2019). Political Secularism and Muslim Integration in the West: Assessing the Effects of the French Headscarf Ban. *American Political Science Review, Forthcoming*.

Abramitzky, R. and L. Boustan (2017). Immigration in American Economic History. *Journal of Economic Literature 55*(4), 1311–45.

Abramitzky, R., L. Boustan, and K. Eriksson (2020). Do Immigrants Assimilate More Slowly Today Than in the Past? *American Economic Review: Insights 2*(1), 125–141.

Achard, P. (2020). Ethnic Enclaves and Cultural Assimilation. Technical report.

Adda, J., P. Pinotti, and G. Tura (2020). There's More to Marriage than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations. *CEPR Discussion Paper No. DP14432*.

Ahn, S. Y. (2018). Matching Across Markets: Theory and Evidence on Cross-Border Marriage. *Working Paper*.

Alesina, A., P. Giuliano, and N. Nunn (2013). On The Origins of Gender Roles: Women and The Plough. *Quarterly Journal of Economics 128*(2), 469–530.

Algan, Y., A. Bisin, A. Manning, and T. Verdier (2012). *Cultural Integration of Immigrants in Europe*. Oxford University Press.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2020). Bride Price and Female Education. *Journal of Political Economy 128*(2), 591–641.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2019). Unity in Diversity? How Intergroup Contact Can Foster Nation Building. *American Economic Review 109*(11), 3978–4025.

Becker, G. S. (1960). *An Economic Analysis of Fertility*. Columbia University Press.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political economy 81*(4), 813–846.

Becker, G. S. (1974). A theory of marriage: Part ii. *Journal of political Economy 82*(2, Part 2), S11–S26.

Bisin, A., G. Topa, and T. Verdier (2004). Religious Intermarriage and Socialization in the United States. *Journal of Political Economy 112*(3), 615–664.

Bisin, A. and T. Verdier (2000). Beyond The Melting Pot: Cultural Transmission, Marriage, And The Evolution of Ethnic and Religious Traits. *Quarterly Journal of Economics 115*(3), 955–988.

Bisin, A. and T. Verdier (2001). The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory 97*(2), 298–319.

Bisin, A. and T. Verdier (2011). The Economics of Cultural Transmission and Socialization. *Handbook of Social Economics*, 339–416.

Bisin, A. and G. Zanella (2017). Time-consistent Immigration Policy under Economic and Cultural Externalities. *Economic Policy 32*(91), 415–446.

Boas, F. (1928). *Anthropology and modern life*. Routledge.

Borjas, G. J. (2003). The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *Quarterly Journal of Economics 118*(4), 1335–1374.

Borjas, G. J. (2014). *Immigration Economics*. Harvard University Press.

Browning, M., P.-A. Chiappori, and Y. Weiss (2014). *Economics of the Family*. Cambridge University Press.

Brunello, G. and D. Checchi (2007). Does School Tracking Affect Equality of Opportunity? New International Evidence. *Economic Policy 22*(52), 782–861.

Card, D. (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial and Labor Relations Review 43*(2), 245–257.

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

Card, D. and G. Peri (2016). Immigration Economics by George J. Borjas: A Review Essay. *Journal of Economic Literature 54*(4), 1333–49.

Carlana, M., E. La Ferrara, and P. Pinotti (2021). Goals and Gaps: Educational Careers of Immigrant Children. *Econometrica*.

Casey, T. and C. Dustmann (2008). Intergenerational Transmission of Language Capital and Economic Outcomes. *Journal of Human Resources 43*(3), 660–687.

Cavalli-Sforza, L. L. and M. W. Feldman (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press.

Chiappori, P.-A. (1988). Rational Household Labor Supply. *Econometrica 56*(1), 63–90.

Chiappori, P.-A. (1992). Collective Labor Supply and Welfare. *Journal of Political Economy 100*(3), 437–467.

Chiappori, P.-A. (2020). The Theory and Empirics of the Marriage Market. *Annual Review of Economics 12*.

Chiappori, P.-A., M. C. Dias, and C. Meghir (2018). The Marriage Market, Labor Supply, and Education Choice. *Journal of Political Economy 126*(S1), S26–S72.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002). Marriage Market, Divorce Legislation, and Household Labor Supply. *Journal of Political Economy 110*(1), 37–72.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009). Investment in Schooling and the Marriage Market. *American Economic Review 99*(5), 1689–1713.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2015). The becker–coase theorem reconsidered. *Journal of Demographic Economics 81*(02), 157–177.

Chiappori, P.-A. and B. Salanié (2016). The Econometrics of Matching Models. *Journal of Economic Literature 54*(3), 832–861.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017). Partner Choice, Investment in Children, and the Marital College Premium. *American Economic Review 107*(8), 2109–67.

Choo, E. (2015). Dynamic Marriage Matching: An Empirical Framework. *Econometrica 83*(4), 1373–1423.

Choo, E. and A. Siow (2006). Who Marries Whom and Why. *Journal of Political Economy 114*(1), 175–201.

Ciscato, E. and S. Weber (2016). Birds of a feather flock together: An analysis of marriage patterns in the united states. Technical report.

Clots-Figueras, I. and P. Masella (2013). Education, Language and Identity. *The Economic Journal 123*(570), 333–357.

Corno, L., N. Hildebrandt, and A. Voena (2020). Age of Marriage, Weather Shocks, and the Direction of Marriage Payments. *Econometrica 88*(3), 879–915.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household Choices and Child Development. *The Review of Economic Studies 81*(1), 137–185.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2012). The Intergenerational Transmission of Risk and Trust Attitudes. *The Review of Economic Studies 79*(2), 645–677.

Dupuy, A. and A. Galichon (2014). Personality Traits and The Marriage Market. *Journal of Political Economy 122*(6), 1271–1319.

Dustmann, C. and F. Fabbri (2003). Language Proficiency and Labour Market Performance of Immigrants in the UK. *The Economic Journal 113*(489), 695–717.

Dustmann, C., S. Machin, and U. Schönberg (2010). Ethnicity and Educational Achievement in Compulsory Schooling. *The Economic Journal 120*(546), 272–297.

Dustmann, C., U. Schönberg, and J. Stuhler (2016). The Impact of Immigration: Why Do Studies Reach Such Different Results? *Journal of Economic Perspectives 30*(4), 31–56.

Dustmann, C., U. Schönberg, and J. Stuhler (2017). Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment. *The Quarterly Journal of Economics 132*(1), 435–483.

Dustmann, C., K. Vasiljeva, and A. Piil Damm (2019). Refugee Migration and Electoral Outcomes. *The Review of Economic Studies 86*(5), 2035–2091.

Fernández, R. (2011). Does Culture Matter? In *Handbook of Social Economics*, Volume 1, pp. 481–510. Elsevier.

Fernández, R. and A. Fogli (2006). Fertility: The Role of Culture and Family Experience. *Journal of the European Economic Association 4*(2-3), 552–561.

Fernández, R. and A. Fogli (2009). Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics 1*(1), 146–177.

Fouka, V. (2020). Backlash: The Unintended Effects of Language Prohibition in US Schools After World War I. *The Review of Economic Studies 87*(1), 204–239.

Fouka, V. (2024, 04). Assimilation in Historical Political Economy. In *The Oxford Handbook of Historical Political Economy*. Oxford University Press.

Furtado, D. and S. J. Trejo (2013). Interethnic Marriages and their Economic Effects. *International Handbook on the Economics of Migration*, 276.

Galichon, A. (2018). *Optimal Transport Methods in Economics*. Princeton University Press.

Galichon, A., S. D. Kominers, and S. Weber (2019). Costly concessions: An empirical framework for matching with imperfectly transferable utility. *Journal of Political Economy 127*(6), 2875–2925.

Galichon, A. and B. Salanié (2017). The Econometrics and Some Properties of Separable Matching Models. *American Economic Review 107*(5), 251–55.

Galichon, A. and B. Salanié (2021). Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *The Review of Economic Studies*.

Gayle, G.-L. and A. Shephard (2019). Optimal Taxation, Marriage, Home Production, and Family Labor Supply. *Econometrica 87*(1), 291–326.

Ginsburgh, V. and S. Weber (2011). *How Many Languages Do We Need?: The Economics of Linguistic Diversity*. Princeton University Press.

Giuliano, P. (2007). Living Arrangements in Western Europe: Does Cultural Origin Matter? *Journal of the European Economic Association 5*(5), 927–952.

Gordon, M. (1964). *Assimilation in American Life*. Oxford University Press.

Guirkinger, C., J.-P. Platteau, and Z. Wahhaj (2019). Behind the Veil of Cultural Persistence: Marriage and Divorce in a Migrant Community. *Working Paper*.

Hwang, Y. G. (2019). An Estimable General-Equilibrium Structural Model of Immigrants' Neighborhood Sorting and Social Integration. *Working Paper*.

Lévi-Strauss, C. (1949). *Les structures élémentaires de la parenté*, Volume 2. Walter de Gruyter GmbH & Co KG.

Low, C. (2024). The Human Capital–Reproductive Capital Trade-Off in Marriage Market Matching. *Journal of Political Economy 132*(2), 000–000.

Lundberg, S. and R. A. Pollak (1993). Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy 101*(6), 988–1010.

Manning, A. and S. Roy (2010). Culture Clash or Culture Club? National Identity in Britain. *Economic Journal 120*(542), F72–F100.

Mayer, E. (2013). *Love & tradition: marriage between Jews and Christians*. Springer.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. *Frontiers in Econometrics*, 105–142.

Meng, X. and R. G. Gregory (2005). Intermarriage and the Economic Assimilation of Immigrants. *Journal of Labor Economics 23*(1), 135–174.

Riesman, P. (1992). *First find your child a good mother: The construction of self in two African communities*. Rutgers University Press.

Salari, M. (2020). Culture and Heritage Language: A Study of Female Labor Force Participation. *Review of Economics of the Household*, 1–22.

Schwartz, C. R. (2013). Trends and Variation in Assortative Mating: Causes and Consequences. *Annual Review of Sociology 39*, 451–470.

Siow, A. (2015). Testing becker's theory of positive assortative matching. *Journal of Labor Economics 33*(2), 409–441.

Smith, R. C. (1996). *Two Cultures One Marriage: Premarital Counseling for Mixed Marriages*. Andrews University Press.

Spolaore, E. and R. Wacziarg (2009). The Diffusion of Development. *Quarterly Journal of Economics 124*(2), 469–529.

Spolaore, E. and R. Wacziarg (2016). Ancestry, Language and Culture. In *The Palgrave Handbook of Economics and Language*, pp. 174–211. Springer.

Storn, R. and K. Price (1997). Differential Evolution–A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization 11*(4), 341–359.

Tabellini, M. (2020). Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration. *Review of Economic Studies 87*(1), 454–486.

Tura, G. (2020). For Better and For Worse. The Dissolution of Interethnic Marriages in Italy. Technical report.

Vieira, C. C., S. Lohmann, E. Zagheni, P. O. Vaz de Melo, F. Benevenuto, and F. N. Ribeiro (2022). The Interplay of Migration and Cultural Similarity Between Countries: Evidence from Facebook Data on Food and Drink Interests. *Plos one 17*(2), e0262947.

Voena, A. (2015). Yours, Mine and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples? *American Economic Review 105*(8), 2295–2332.

For Online Publication

A Data and Sample Construction

We obtained restricted accessto administrative Italian data at the individual level from ISTAT through its ADELE Laboratory.[58] In what follows, we start describing our data sources and variables of interest; passing then onto a discussion of the sample construction, and finally to the computation of empirical moments.

A.1 Marriages, Fertility, Separation and Singles

Marriage. Our empirical analysis takes advantage of administrative individual-level data from the Italian Statistical Institute (ISTAT, ADELE Laboratory), covering the universe of marriages celebrated in Italy from 1995 to 2012 (over 4 million marriages). We exploit marriage records from municipal vital statistics registries to recover matching patterns by ethnic group of the spouses. Marriage records contain the universe of marriages celebrated each year in Italy from 1995 to 2012. They provide information on the main socio-demographic characteristics of the spouses. They are collected through the ISTAT model compiled by the Registrar of the City Civil State in which the marriage took place. For each marriage, the section dedicated to the wedding reports: the date of marriage, the type of ceremony (religious or civil), the municipality of the ceremony, and the choice of the property regime by the spouses (community or separation property). The information provided for each spouse includes: date of birth, municipality of birth, municipality of residence at the time of marriage, the place of future residence of the spouses, the previous marital status, the education level, the employment status, and for immigrant individuals the nationality and the country of origin. In order to account for the out-migration selection of families, the sample is restricted to marriages where at least one spouse is resident in Italy at the time of the marriage.

From these data, we recover the distribution of marriages, by cultural-ethnic group and education of spouses and by region; see Appendix Table C.2.[59]

[58] Requests for accessing the data for research purposes should be addressed to ISTAT through an open application procedure. Authorized researchers can access and use the data from work stations located in secure rooms within the ISTAT offices. The output of the analysis is made available upon inspection by ADELE officers in compliance with the laws on the protection of statistical confidentiality and personal data. For further information, visit https://www.istat.it/it/informazioni-e-servizi/per-i-ricercatori/laboratorio-adele.

[59] These data account only for legal marriages and only when celebrated in Italy. The increase in cohabiting couples is a very recent phenomenon. In 2011, the share of cohabiting native couples was still only about 9% in Italy - compared to the averages in the EU (14.5%) and the OECD (16.7%). Also, cohabitations are less

Fertility. Fertility rates come from municipality births registries, which contain the universe of individual birth records of residents in the municipality of enrolment, for each year from 1990 to 2012. Individual birth records include socio-demographic variables of interest such as gender, date and province of birth, citizenship and parental information regarding their date of birth, citizenship and marital status.

Separation. Separation data come from the registries of civil court chancelleries and cover the universe of legal separations registered in Italy, covering the period 1995-2012.[60] We focus on separation rates, which better represent marital dissolution decisions in the Italian context compared to divorces, for two main reasons. First, separation is the juridical act that launches the divorce proceedings. With Law 74/1987 and until 2015, a minimum period of 3 years of legal separation was required before eventually submitting a divorce request. Second, on average only 65% of separations are followed by a divorce, which implies that divorce choices significantly underestimate marital dissolution behaviours. The data allow us to analyse various aspects of the marital dissolution phenomenon. We investigate, in particular, the custody assignment of children. [61]

Single Individuals. We derive the population vectors by ethnic group, gender and marital status from individual Italian Census data of 2001 and 2011. We select adult individuals, hence the age range we focus on is of more than 18 years old. Census data classify the marital status of an individual as: never married, at present married, separated *de facto*, legally separated, divorced or widowed. We consider an individual available in the case that she/he is never married, legally separated, divorced or widowed. We also discard institutional households, corresponding to correctional institutions, but also military and mental care facilities. We take into account potential measurement error concerns due to truncation

relevant for our study since they are typically less stable and hence less motivated by fertility and children socialization; see (Lundberg and Pollak, 1993; Chiappori et al., 2017). Also, marriages celebrated out of the Italian marriage market are arguably less motivated by cultural integration issues.

[60] For our investigation period, registries of civil court chancelleries constitute the unique source for separations data, while starting from December 2014 (in application of Law n. 162/2014) consensual separation proceedings can be submitted to the civic registrar. This rules out potential sample selection concerns.

[61] In our model, we introduce an asymmetry between spouses in the probability of child custody assignment upon dissolution, independently from the ethnic-groups h, j. From separation proceedings data, we calculate that the mother is given *effective* custody of children in 88% of the cases. We uncover some significant differences in custody assignment conditional on mother and father migrant status, but we abstract from incorporating them in the model for the sake of simplicity. Specifically, foreign mothers married with a native husband are less likely to obtain their child's custody by 1.4 (3.4) p.p. compared to native mothers, upon separation (divorce). On the contrary, native mothers are more likely to obtain their children custody by 5.1 (6.9) p.p. following a separation (divorce) if married with a foreign husband.

of unmatched population vectors, we follow Chiappori et al. (2017). Specifically, to account for the possibility that single individuals might marry in the near future, we restrict the set of single individuals to single men and women after their marriageable age, defined as the 90th percentile of the age at first marriage distribution for men and women, respectively. In our data, single rates increase quite symmetrically for all ethnic groups, from 2001 to 2011, the overall Spearman rank correlation test is as high as 0.88, and equal to 0.57 and 0.98 for available adult men and adult women, in turn, suggesting that the ethnic-group rank order remains stable over the period, especially for women.

Socialization. Socialization data come from the *Condition and Social Integration of Foreign Nationals* Survey, conducted in 2011 and 2012 in all Italian regions on a sample of 9,600 families. The survey targeted foreign residents in Italy and it was conducted at the household level to provide socio-demographic information about all family members, for a total sample of 25,356 respondents.[62] The final questionnaire was translated in 10 different languages to overcome potential language barriers and to reduce attrition. The actual survey was conducted through direct interviews supported by the CAPI (Computer Assisted Personal Interview) system to ease the development of the whole questionnaire.[63] In each selected household, all members were interviewed, both foreign-born and natives.

We proxy the cultural-ethnic transmission with language socialization. In particular, the socialization measure we construct for our analysis is based on the *language spoken at home* by children and young adults (less than 25 years old), living with their parents at the time of the interview: an individual is socialized to the Italian language if he/she declares to speak Italian within the family; otherwise, we assume he is socialized to his mother language, defined as idiom acquired during the preschool period of childhood.[64]

We exclude from our analysis, respondents who are single and families without children, at the time of the interview. For our analysis, we consider children and young adult of less than 25 years old, living with their parents at the time of the interview. The final sample consists of 8,007 individuals belonging to about 5,000 families, 86.7% of these families are married while the remaining are either separated or divorced. We consider the sample representative for the study of immigrant linguistic integration by ethnic group in each region of residence.

[62] The survey follows a pivotal survey conducted in 5 sampled regions on a sample of 250 families with at least one foreign member. The pivotal survey was particularly useful in the definition and evaluation of the questionnaire, which also requires the participation of sociologists and cultural mediators.

[63] Examples of the questionnaire and invitation letter are available at http://www.istat.it/it/archivio.

[64] The three questions we exploit are framed in the survey in the following way. Language spoken at home: *In Italy, in your family, do you speak more often Italian or another language?*. Mother tongue (main): *What language did you speak when you were young, before going to school?*. Mother tongue (secondary): *In addition to this, did you also speak another language when you were young and which one?*

We construct our measure of socialization based on the language spoken at home. The survey also provides questions to evaluate the level of Italian language proficiency and we check individual self-declared responses on language spoken.

A.2 Dataset Construction and Empirical Moments

The empirical estimation is based on a unique dataset that links households information across different sources. We matched marriage, birth and separation records on the exact date of marriage and spouses' exact date and place of birth (Italian province for natives and country of origin for foreigners), which are reported in all registries. In the birth records matching, the combination of these characteristics allows for an exact one-to-one matching for 98.8% of marriages, while in the separation matching, we match exactly the 99.5% of marriages, and we discard the remaining fraction. Such low percentages suggest that marriages can be uniquely identified through the set of time-invariant characteristics listed above. The final sample of 4 mil marriages corresponds to 92.58% of the universe of marriages celebrated in Italy during the time interval 1995-2012. In the final dataset the fertility rate corresponds to 69.56% with an average of 1.54 children per family. Of all marriages, 7% end up in separation in the first years of the marital union.

From this final sample, we recover the following empirical moments. The marital utility net of the outside options of singlehood $\hat{U}_{hj}$ for the household of type hj is identified from equation (1), exploiting the number of hj marriages formed in each region r, μ_{hj}, and the number of unmatched men of type h and women of type j for each region r, $\mu_{h.}, \mu_{.j}$.

Fertility rates $\hat{N}_{hj}$ and separations rates $\hat{\pi}_{hj}$ for each household type hj and for all regions r are computed as follows:

$$\hat{N}_{hj} = \frac{1}{\mu_{hj}} \sum_{m=1}^{\mu_{hj}} N_{hj},$$

$$\hat{\pi}_{hj} = \frac{1}{\mu_{hj}} \sum_{b=1}^{\mu_{hj}} D_{hj},$$

with N_{hj} the number of children born from within a hj household, and D_{hj} is a dummy equal to one if the hj marriage end up in separation during the investigation period.

We compute the vector of socialization frequencies $\hat{P}^k_{hj}(d)$ for all h, j and k, conditional on being married, $d = 0$, and for all regions r, as follow:

$$\hat{P}^k_{hj}(d=0) = \frac{1}{M_{hj}} \sum_{b=1}^{M_{hj}} S^k_{hj}.$$

with M_{hj} being the number of children and young adults of less than 25 years old belonging to the hj household, and speaking language S^k. Due to data limitations in the number of divorced households per type of family and region, in the estimation we exploit only socialization moments for married families.

B Education in Italy and INVALSI data

We obtained administrative student-level data on standardized reading and math proficiency tests, as well as related survey data from the National Institute for the Evaluation of the Italian Education System (INVALSI).[65] In what follows, we start describing the Italian educational system, our data sources and variables of interest; passing then onto a discussion of our sample construction.

B.1 Italian educational setting

In Italy, pupils normally enter formal schooling at the age of 6, and education is compulsory for 10 years. The Italian educational system is organized in five grades of elementary school, three grades of middle school, and five grades of high school. For each school (elementary, middle and high school), students are assigned to classes and take all their subjects within the same class and with the same peers. In elementary and middle school, the educational curriculum is the same for all pupils and the subjects studied are the same.

High school is divided into different tracks (academic, technical and vocational) and students freely self-select into three different tracks. The three tracks have the same duration, but differ widely in terms of curriculum, difficulty, and prestige. While in principle, access to university is also possible from some schools within the vocational track; in practice, academic and technical schools offer much better educational and career prospects. Following Carlana et al. (2021), we define academic and technical schools as *high-track* schools, and we refer to vocational schools as the *low-track* ones. This early stratification in high school tracks ultimately have long-term implications for access to college and occupational careers (Brunello and Checchi, 2007; Carlana et al., 2021).

B.2 INVALSI tests and survey data

INVALSI tests. Every year, starting from 2010, INVALSI administers standardized tests in reading and math to the entire population of Italian students. Tests are administered at various points of students' careers, specifically at the end of grades 2, 5, 6, 8, and 10. The INVALSI test is identical for all students in a given grade, it is blindly marked by an external evaluator following a precise evaluation scheme, hence students' results are objective and fully comparable across schools in Italy. We exploit standardized test scores, with zero mean and standard deviation of one. The test consists of multiple-choice and open-

[65] Requests for accessing the data for research purposes should be addressed to INVALSI through an open application procedure. For further information, visit https://invalsi-serviziostatistico.cineca.it/.

ended questions, where the exact structure varies by grade.[66] **INVALSI questionnaire.** Besides test scores, INVALSI data provides rich information for each student, including demographic characteristics such as year and quarter of birth, gender, citizenship acquisition, grade retention, and family background characteristics such as parents' education, migration history, macro-area of origin, employment status, and some measures of socio-economic status. Crucially for us, INVALSI also collects additional data from a students' questionnaire, including a specific question on the main language spoken at home by students, similarly to our main analysis. The questionnaire is administered only in grade 5 (the last year of elementary school) and in grade 10 (second year of high school).

B.3 Sample selection

In our analysis, we focus on students enrolled in grade 5 between school years 2012-13 and 2018-19. For these pupils the language spoken at home is likely a choice of the parents. Thanks to a unique student identifier, we are able to follow students over time and match the scores and information of students in grade 5 with their educational careers choices in grade 10. Because of the data collection scheme, we are able to track only two cohorts of students, i.e., the students enrolled in grade 5 in school years 2012-2013 and 2013-14 that we observe in grade 10 in school years 2017-2018 and 2018-19. Finally, our sample includes all students with at least one immigrant parent.

B.4 Probability of being high educated

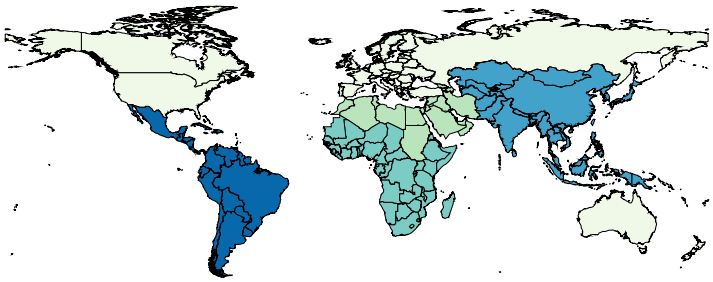
We exploit INVALSI data to recover the probability that children become high educated, i.e., enroll into *high-track* schools compared to vocational schools, conditional on family type. We identify native and immigrant parents thanks to country of origin, and among immigrants we distinguish immigrants coming from European-EU15 and Other European countries. All other immigrant minorities are grouped together.

[66] The tests are designed to align with those administered by the OECD Programme for International Student Assessment (PISA).

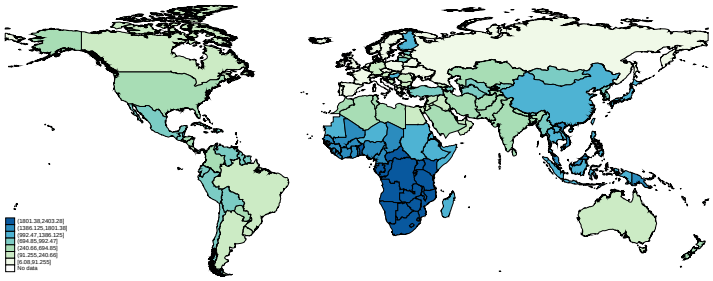
C Additional Figures and Tables

Figure C.1: Ethnic-Group Classification and Cultural Distance wrt Italy

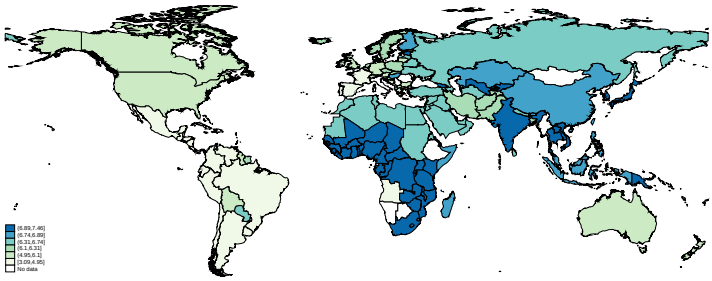
(a) Our Cultural-Ethnic Group Classification



(b) Genetic Distance to Italy



(c) Linguistic Distance to Italy



Notes: This figure shows our classification of countries in cultural-ethnic groups (panel a) and plots the cultural distance of each country towards Italy as proxied by genetic (panel b) and ethnolinguistic distance (panel c). Data for genetic and ethnolinguistic distance are available thanks to Spolaore and Wacziarg (2016). The classification of immigrants' countries of origin by cultural-ethnic group is reported in Table C.1. In particular, our classification parallels the heterogeneity in genetic distance within Africa, between the Arabic countries in the North and Sub-Saharan countries, as well as the within Asia divide between Middle-East and East Asia countries.

Table C.1: Cultural-Ethnic Group Classification of Migrants' Countries of Origin

Cultural-Ethnic Group	(%)	Countries
Europe-EU15, i^E	4.57	Austria, Belgium, Denmark, France, Finland, Germany, Greece, Ireland, Luxembourg, Netherlands, Portugal, United Kingdom, Spain, Sweden
Other Europe, i^O	46.29	Albania, Andorra, Belarus, Bosnia and Herzegovina, Bulgaria, Cyprus, Croatia, Czech Republic, Estonia, Hungary, Iceland, Isle of Man, Liechtenstein, Latvia, Lithuania, Kosovo, Macedonia (FYROM), Malta, Poland, Republic of Moldova, Monaco, Norway, Russian Federation, San Marino, Vatican City State, Serbia and Montenegro, Romania, Switzerland, Slovakia, Slovenia, Turkey, Ukraine, Vatican City State, United States, Canada
North Africa-Middle East, i^M	17.15	Algeria, Egypt, Libyan Arab Jamahiriya, Marocco, Tunisia, Afghanistan, Saudi Arabia, Armenia, Azerbaijan, United Arab Emirates, Islamic Republic Of Iran, Iraq, Israel, Kazakhstan, Kyrgyzstan, Kuwait, Lebanon, Qatar, Syrian Arab Republic, Palestinian Territory, Turkmenistan, Uzbekistan
Sub-Saharan Africa, i^A	7.33	Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo, The Democratic Republic of Congo, Cote D'Ivoire, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Dijbouti, Guinea, Guinea-Bisseau, Equatorial Guinea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Swaziland, United Republic of Tanzania, Togo, Uganda, Zambia, Zimbabwe
East Asia, i^S	16.47	Brunei Darussalam, Cambodia, China, Democratic People's Replica of Korea, Republic of Korea, Philippines, Japan, Jordan, Indonesia, Lao Pepople's Democratic Republic, Malaysia, Mongolia, Myanmar, Singapore, Taiwan, Thailand, East Timor, Vietnam, Australia, Fiji, Kiribati, Marshall Islands, Federated States of Micronesia, Nauru, New Zealand, Palau, Papua New Guinea, Solomon Islands, Samoa, Tonga, Tuvalu, Vanuatu, Bahrain, Bangladesh, Bhutan, Georgia, India, Maldives, Nepal, Oman, Pakistan, Sri Lanka, Tajikistan, Yemen
Latin America, i^L	8.2	Antigua and Barbuda, Argentina, Bahamas, Barbados, Belize, Plurinational State of Bolivia, Brazil, Costa Rica, Cuba, Chile, Colombia, Dominica, Dominican Republic, Ecuador, El Salvador, Jamaica, Grenada, Guatemala, Guyana, Haiti, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and The Grenadines, Suriname, Trinidad and Tobago, Uruguay, Venezuela

Notes: This table reports our classification of foreign countries by cultural-ethnic group.

Table C.2: Marriage Distribution by Spouses Cultural-Ethnic Group and Education

		Female attributes:														
		Italian		Europe-EU15		Other Europe		Middle East		Sub-Sah.Africa		East Asia		Latin America		Total
		L	H	L	H	L	H	L	H	L	H	L	H	L	H	
Male attributes:																
Italian	L	1123913	567812	15042	7735	54593	38289	5065	2166	4555	1715	4071	2710	21021	12749	1861436
	H	282302	1735427	3972	23354	13605	57670	1137	3425	1583	3219	1607	7049	6927	22644	2163921
Europe-EU15	L	10986	5973	853	352	746	444	39	14	54	17	58	30	245	138	19949
	H	3005	19713	216	1643	155	930	19	40	29	86	33	151	86	328	26434
Other Europe	L	12999	6844	360	177	10060	3227	77	34	40	17	67	49	329	199	34479
	H	4666	18888	105	489	1654	9182	22	71	32	76	40	153	120	471	35969
Middle-East	L	9076	2865	175	86	1084	479	2003	466	40	12	28	10	246	120	16690
	H	3485	5995	53	212	269	1063	509	1151	13	36	16	53	73	348	13276
Sub-Sahara Africa	L	1157	1012	44	31	83	76	15	1	4418	723	3	3	33	20	7619
	H	756	2886	22	112	33	177	5	18	1401	3313	5	6	7	51	8792
East Asia	L	1091	525	31	13	126	58	14	2	6	3	6098	912	28	17	8924
	H	448	1932	6	53	45	157	-	12	3	8	644	2106	8	43	5465
Latin America	L	3018	1863	61	33	224	147	11	1	7	-	8	2	4935	1746	12056
	H	1435	6565	29	159	85	436	1	10	7	7	8	21	1142	6368	16273
Total		1458337	2378300	20969	34449	82762	112335	8917	7411	12188	9232	12686	13255	35200	45242	4231283

Notes: This table reports the bivariate marriage distribution by cultural-ethnic group and education of spouses (absolute numbers). We consider seven cultural-ethnic groups: Italian, European-EU15, Other European, North African and Middle-Eastern, Sub-Saharan African, East Asian, and Latin American. Regarding education, we distinguish two groups: individuals with High H (more than high school) and Low L (no high school) education.

Table C.3: Gains from Marriage by Spouses Cultural-Ethnic Group and Education

		Female attributes:													
		Italian		Europe-EU15		Other Europe		Middle East		Sub-Sah.Africa		East Asia		Latin America	
		L	H	L	H	L	H	L	H	L	H	L	H	L	H
Male attributes:															
Italian	L	-0,55	-2,85	-4,92	-6,93	-3,53	-5,03	-6,95	-8,44	-7,01	-8,82	-7,65	-8,44	-4,42	-6,15
	H	-3,82	-1,19	-8,06	-5,28	-6,89	-4,70	-10,36	-8,14	-9,80	-8,11	-10,17	-7,04	-7,35	-5,35
Europe-EU15	L	-5,30	-7,40	-6,45	-9,16	-7,67	-9,31	-10,70	-12,88	-10,76	-12,38	-10,06	-11,52	-8,73	-10,87
	H	-8,05	-5,32	-9,27	-6,31	-11,19	-8,18	-12,22	-11,01	-12,36	-10,11	-11,57	-9,25	-11,15	-9,32
Other Europe	L	-6,05	-8,27	-9,15	-11,41	-4,36	-7,44	-11,62	-12,99	-12,63	-14,03	-12,44	-12,56	-9,46	-11,06
	H	-8,38	-6,39	-11,77	-9,33	-8,09	-5,40	-13,85	-11,74	-13,37	-11,21	-13,10	-10,55	-11,62	-9,34
North Africa-Middle-East	L	-6,60	-9,89	-9,35	-11,93	-7,82	-10,19	-5,75	-8,37	-11,40	-14,19	-12,76	-13,97	-9,67	-11,27
	H	-7,83	-7,63	-11,49	-9,59	-10,37	-8,10	-7,58	-5,63	-13,07	-11,84	-12,82	-10,56	-11,72	-9,23
Sub-Saharan Africa	L	-9,36	-10,91	-10,94	-12,76	-11,25	-12,26	-12,03	-16,32	-3,19	-6,37	-15,40	-15,25	-11,71	-13,25
	H	-10,13	-8,48	-11,71	-9,60	-12,95	-10,52	-14,05	-11,57	-5,31	-3,22	-13,97	-13,54	-13,95	-11,20
East Asia	L	-10,13	-12,67	-11,59	-14,25	-11,15	-13,63	-12,85	-16,27	-13,28	-14,26	-3,12	-6,63	-12,61	-13,58
	H	-11,60	-9,41	-12,81	-10,82	-12,41	-11,02	-	-12,18	-14,01	-11,81	-6,95	-4,47	-14,11	-11,51
Latin America	L	-7,43	-9,57	-10,56	-12,45	-9,71	-11,28	-12,00	-14,25	-13,60	-	-13,20	-16,07	-3,35	-5,83
	H	-9,36	-7,28	-12,19	-9,93	-11,98	-9,55	-14,60	-12,72	-13,45	-13,33	-13,78	-11,93	-6,67	-3,59

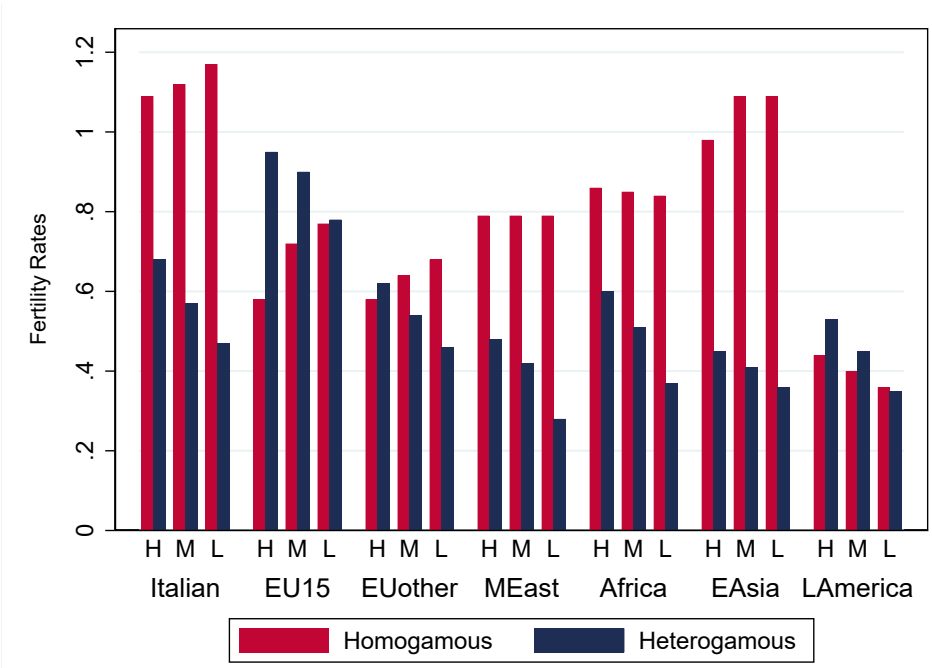
Notes: This table reports average estimates for gains from marriage, computed as in equation (1), across cells defined by spouses' cultural-ethnic group (Italian, European-EU15, Other European, North African and Middle-Eastern, Sub-Saharan African, East Asian and Latin American), education (high - more than high school and low - no high school), and region. We select only adult single men and women (of more than 18 years of age). To account for the possibility that single individuals might marry in the near future, we follow Chiappori et al. (2017); see Appendix A for details.

Table C.4: Correlation of Marriage Outcomes and Cultural Distance Measures

	Panel A. Dep. var: Gains from Marriage			
Genetic Distance	-5.343***			
	(0.324)			
Linguistic Distance		-6.250***		
		(0.139)		
Religious Distance			-6.659***	
			(0.216)	
Cultural distance index, WVS				-8.378***
				(0.455)
Observations	2368	2368	2368	2368
R2	0.437	0.666	0.720	0.598
	Panel B. Dep. var: Fertility rates			
Genetic Distance	-0.452***			
	(0.047)			
Linguistic Distance		-0.572***		
		(0.042)		
Religious Distance			-0.617***	
			(0.053)	
Cultural distance index, WVS				-0.816***
				(0.094)
Observations	2368	2368	2368	2368
R-squared	0.285	0.489	0.541	0.483
	Panel C. Dep. var: "Italian spoken at home"			
Genetic distance	-0.053			
	(0.043)			
Linguistic distance		-0.120		
		(0.293)		
Religious Distance			-0.031	
			(0.068)	
Cultural distance index, WVS				-0.036
				(0.060)
Observations	596	596	596	596
R-squared	0.339	0.335	0.336	0.336

Notes: This table shows OLS estimates of the relationship between marriage outcomes and cultural distance. The outcomes are gains from marriage, computed as in equation (1), in Panel A; fertility and Italian socialization rates in Panel B and C, in turn. We consider four different measures of cultural distance, as explanatory variables, i.e., distance along genetics, language, religion, and attitudes and values from the World Value Survey; we refer to Spolaore and Wacziarg (2009, 2016) for further details. All specifications control for spouses' education and include marriage market (at the regional level) FEs. Robust standard errors clustered at the regional level in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Figure C.2: Fertility Rates by Ethnic Group and Education of Spouses



Notes: This Figure shows average fertility rates by ethnic group and education of spouses. Estimates are reported separately for homogamous (red) and heterogamous (blue) families.

Table C.5: Fertility Rate by Culture and Education

Dep. var:	(1)	(2)	(3)	(4)
	Fertility rate			
Homogamous by culture	0.280***	0.280***	0.280***	0.337***
	(0.016)	(0.017)	(0.017)	(0.023)
Homogamous by education		-0.005		
		(0.009)		
High-high educ			0.026	0.050**
			(0.019)	(0.021)
High and low educ			0.019	0.031*
			(0.013)	(0.015)
Hom by culture × High-high educ				-0.114***
				(0.021)
Hom by culture × High and low educ				-0.057***
				(0.017)
Observations	2456	2456	2456	2456
R-squared	0.296	0.315	0.315	0.318
Marriage market (region) FE	Yes	Yes	Yes	Yes
Spouses' Cultural-ethnic group FE	Yes	Yes	Yes	Yes

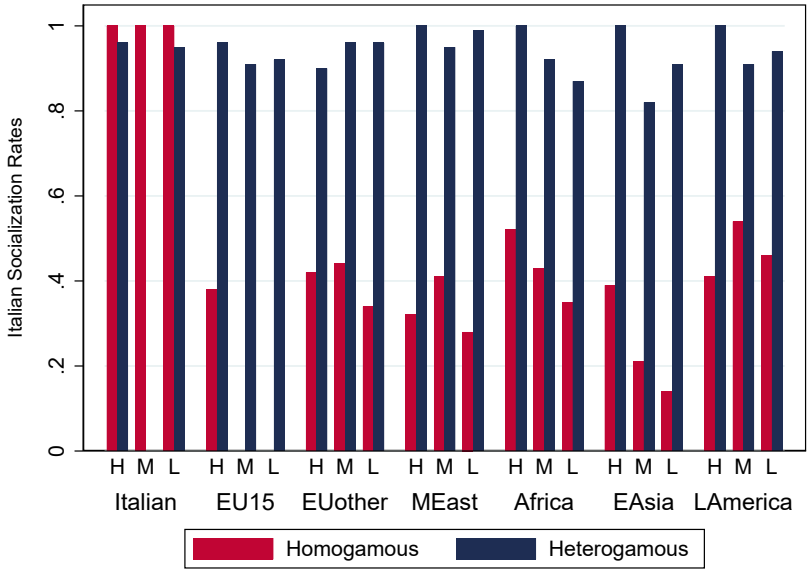
Notes: The dependent variable in the regression is the fertility rate in marriages defined by spouses' cultural ethnic-group and education, and by region. The explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and education (*Homogamous by education*). Columns 3-5 include dummies for marriages by education, dividing marriages of high-educated (*High-high educ*) and mixed education marriages of (*High and low educ*) spouses. Column 5 includes interaction variables. All specifications include marriage market (at the regional level) FEs effects, as well as wife educational level and cultural-ethnic group of origin FEs. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.6: Separation Rate by Culture and Education

Dep. var:	(1)	(2)	(3)	(4)	(5)
	Separation rate				
Homogamous by culture	-0.015***		-0.015***	-0.015***	-0.020***
	(0.003)		(0.003)	(0.003)	(0.006)
Homogamous by education		0.013***	0.013***		
		(0.002)	(0.002)		
High educ (at least one)				-0.020***	-0.022***
				(0.004)	(0.004)
Hom by culture × High educ (at least one)					0.007
					(0.005)
Observations	2456	2456	2456	2456	2456
R-squared	0.090	0.063	0.101	0.104	0.105
Marriage market (region) FE	Yes	Yes	Yes	Yes	Yes
Spouses' Cultural-ethnic group FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable in the regression is the separation rate in marriages defined by spouses' cultural ethnic-group and education, and by region. The explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and by education (*Homogamous by education*). We also include a dummy for marriages with at least one high educated spouses (*High educ, at least one*) and its interaction with the dummy for cultural homogamous marriages (*Hom by culture × High educ*). All specifications include marriage market (at the regional level) fixed effects as well as spouses' cultural-ethnic group of origin fixed effects, and a dummy for wife's high education. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Figure C.3: Italian Socialization Probabilities by Ethnic Group and Education of Spouses



Notes: This figure shows Italian socialization probabilities by ethnic group and education of spouses. The outcome variable "Italian spoken at home" is an indicator for whether the child speaks Italian within the family. Estimates are reported separately for homogamous (red) and heterogamous (blue) families.

Table C.7: Fraction of "Italian spoken at home" by Culture and Education

Dep. var:	(1)	(2)	(3)	(4)
	"Italian spoken at home"			
Homogamous by culture	-0.320***	-0.319***	-0.322***	-0.409***
	(0.023)	(0.022)	(0.023)	(0.031)
Homogamous by education		-0.024		
		(0.022)		
High-high educ			0.044	-0.034
			(0.075)	(0.074)
High and low educ			0.050	-0.050
			(0.058)	(0.071)
Hom by culture × High-high educ				0.068
				(0.057)
Hom by culture × High and low educ				0.141***
				(0.039)
Observations	615	615	615	615
R-squared	0.604	0.608	0.608	0.613
Marriage market (region) FE	Yes	Yes	Yes	Yes
Spouses' Cultural-ethnic group FE	Yes	Yes	Yes	Yes

Notes: The dependent variable in the regression is the fraction of "Italian spoken at home" in marriages defined by spouses' cultural ethnic-group and education, and by region. The explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and education (*Homogamous by education*). Columns 3-5 include dummies for marriages by education, dividing marriages of high-educated (*High-high educ*) and mixed education marriages of (*High and low educ*) spouses. Column 5 includes interaction variables. All specifications include marriage market (at the regional level) FEs effects, as well as wife educational level and cultural-ethnic group of origin FEs. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.8: Fraction of "High educated Children" by Culture and Education

	(1)	(2)	(3)	(4)	(5)
Dep. var:	Fraction of "High educated Children"				
Homogamous by culture	-0.081***		-0.080***	-0.081***	-0.053***
	(0.013)		(0.013)	(0.013)	(0.015)
Homogamous by education		0.013**	0.015**		
		(0.006)	(0.006)		
High educ (at least one)				0.071***	0.079***
				(0.006)	(0.006)
Hom by culture × High educ (at least one)					-0.038***
					(0.004)
Observations	2456	2456	2456	2456	2456
R-squared	0.508	0.301	0.618	0.640	0.641
Marriage market (region) FE	Yes	Yes	Yes	Yes	Yes
Spouses' Cultural-ethnic group FE	Yes	Yes	Yes	Yes	Yes

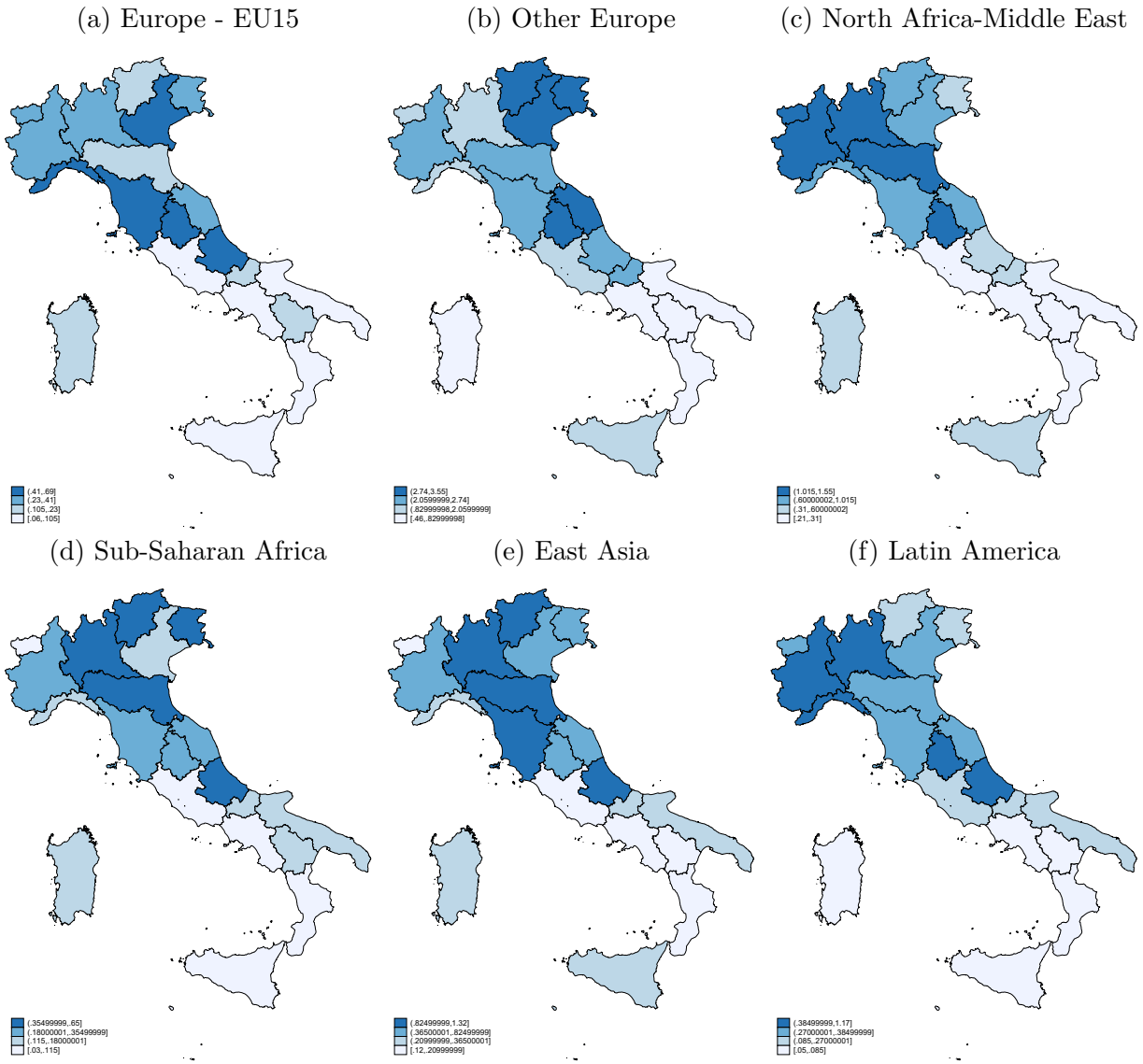
Note: The dependent variable in the regression is the probability of "High educated children" in marriages defined by spouses' cultural ethnic-group and education, and by region.The explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and by education (*Homogamous by education*). We also include a dummy for marriages with at least one high educated spouses (*High educ, at least one*) and its interaction with the dummy for cultural homogamous marriages (*Hom by culture* × *High educ*). All specifications include marriage market (at the regional level) fixed effects as well as spouses' cultural-ethnic group of origin fixed effects, and a dummy for wife's high education. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.9: Separation Rates in Marriages With and Without Children

	Separation Rates			
	Homogamous		Heterogamous	
	$\pi_{hh}\,(n>0)$	$\pi_{hh}\,(n=0)$	$\pi_{hj}\,(n>0)$	$\pi_{hj}\,(n=0)$
Italian	0.054	0.095	0.045	0.097
Europe-EU15	0.024	0.025	0.041	0.061
Other Europe	0.016	0.040	0.039	0.093
North Africa-Middle East	0.023	0.072	0.073	0.127
Sub-Saharan Africa	0.017	0.037	0.063	0.108
East Asia	0.010	0.021	0.040	0.080
Latin America	0.026	0.061	0.053	0.114

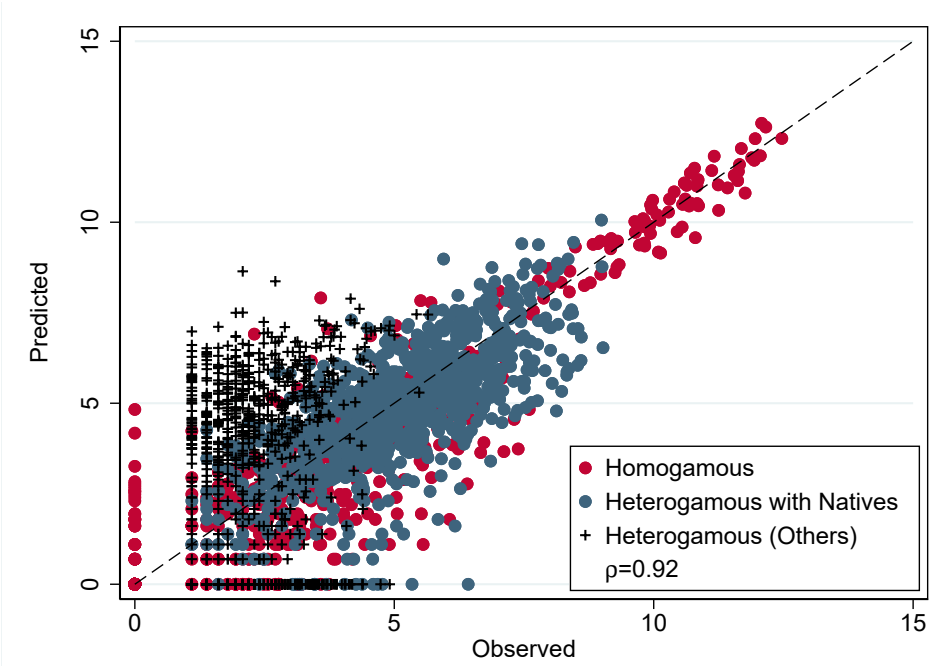
Notes: This table reports the separation rates by ethnic group of spouses in families with and without children, separately for homogamous and heterogamous couples.

Figure C.4: Migrants' Distribution across Regions



Notes: This figure shows the distribution of the immigrant population by cultural-ethnic group and region. Population shares by cultural-ethnic group and region are computed over the total resident population at the regional level. The ethnic group classification is defined in Table C.1. The color classification corresponds to the quartiles of the population distribution.

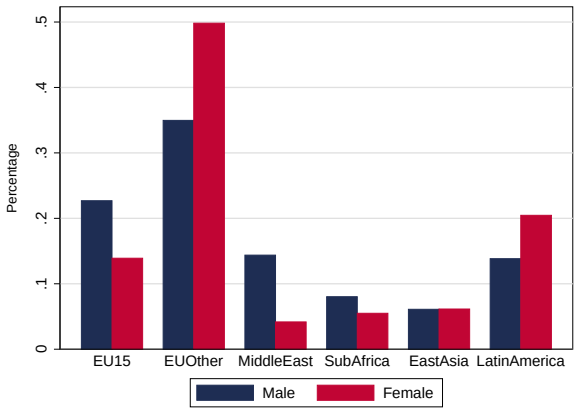
Figure C.5: Fit of the Model - Number of Marriages by Match and Region



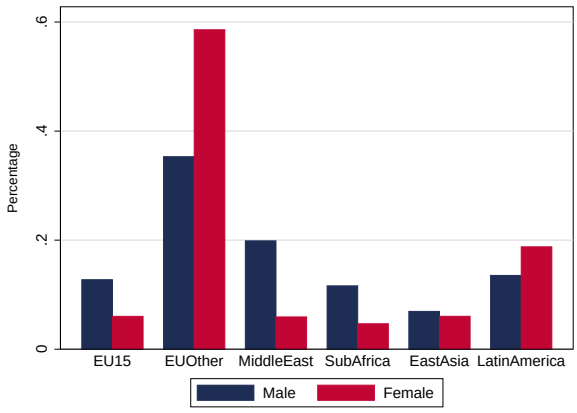
Notes: This figure shows the relationship between the number of marriages observed in the data (in log) and the number of marriages predicted by the model (in log) by region for homogamous families (red), heterogamous families with natives (blue), and all other heterogamous matches (black).

Figure C.6: Distribution of men and women by ethnic group, in the sample and out-of-sample period

(a) Sample period (1995-2012)



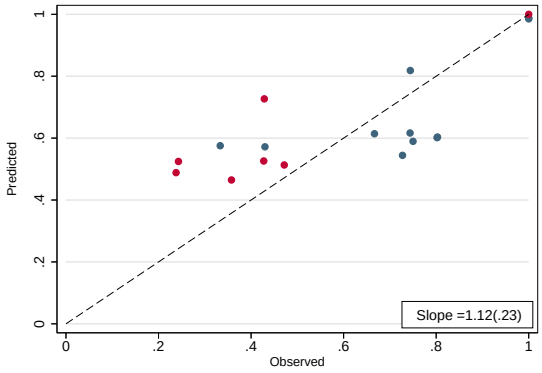
(b) Out-of-sample period (2013-2019)



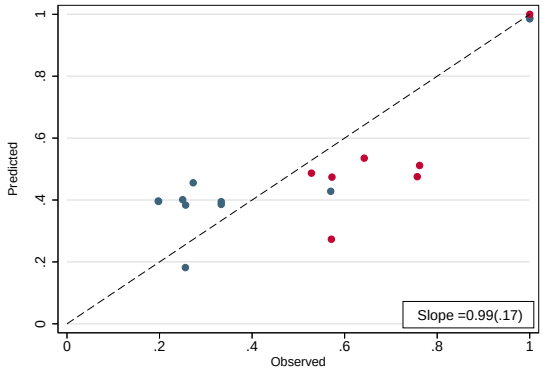
Notes: This figure shows the distribution of male and female population vectors (conditional on getting married) by cultural-ethnic group in the sample (1995-2012) and out-of-sample (2013-2019) period, in turn.

Figure C.7: Model Validation - Non Targeted Socialization Rates for Divorced Couples

(a) Italian Socialization

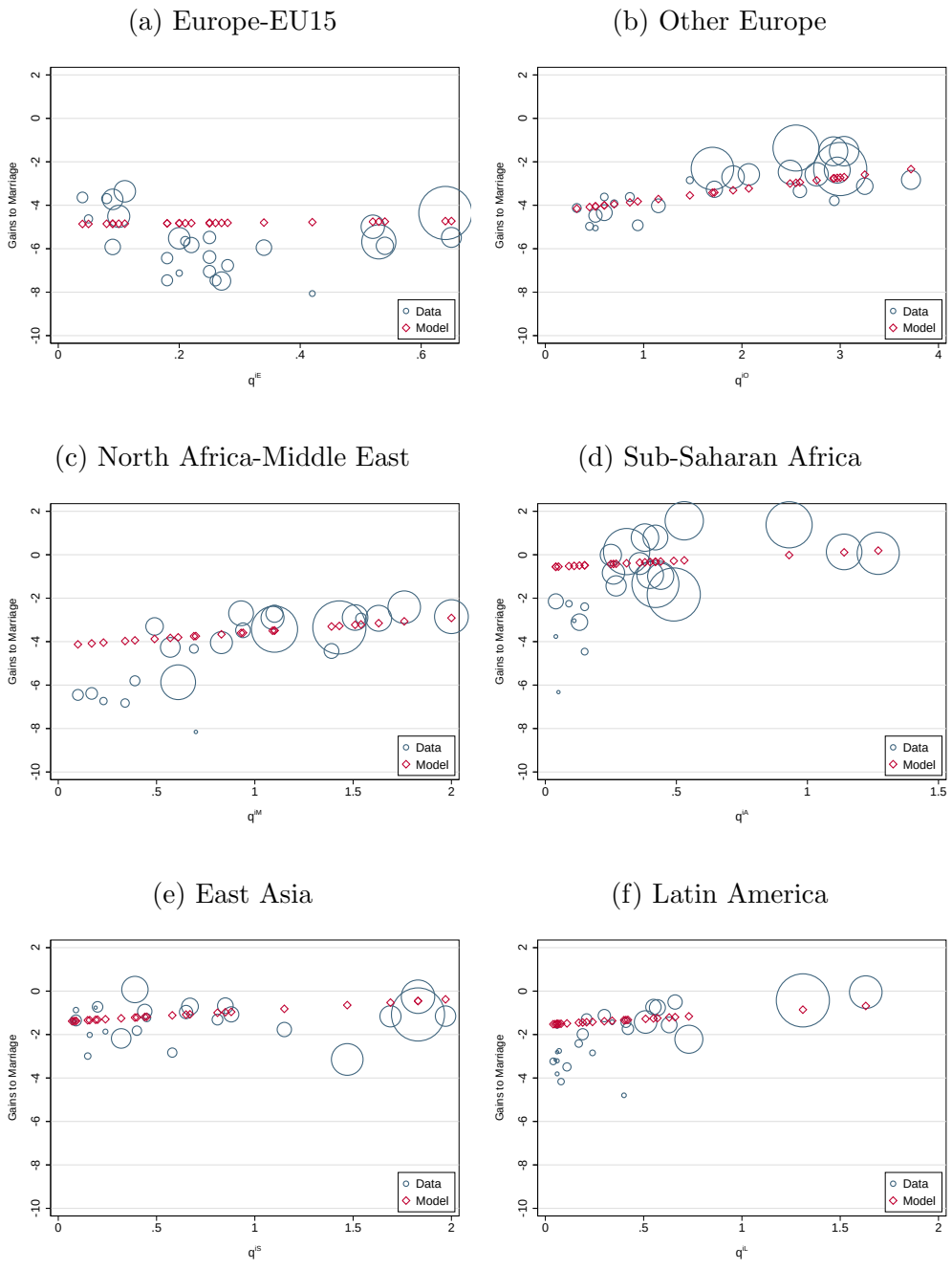


(b) Mother Socialization



Notes: This figure displays the scatterplot of the relationship between the observed and predicted Italian and mother socialization probabilities for the subsample of marriages ending in divorce, separately for homogamous families (red), and heterogamous families with natives (blue). 45-degree dash line also reported.

Figure C.8: Model Validation - Gains from Marriage for Homogeneous Families by Province

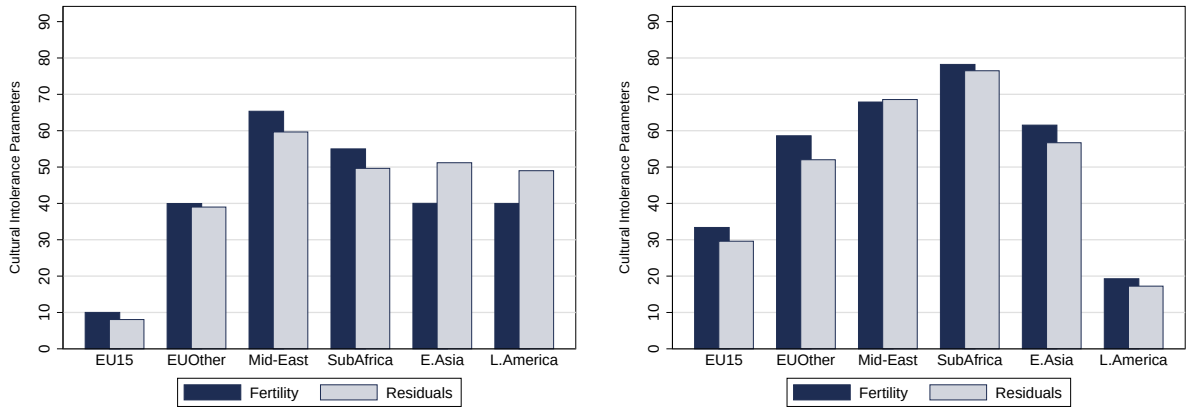


Notes: The figure shows out-of-sample predicted and implied gains from marriage for homogamous families of ethnic group minorities over the corresponding population share, q^i (in percentage), by province of residence (average over the time period). Empirical moments are weighted by the observed number of marriages per province. We select the most representative provinces across northern, central and southern parts of the country. The provinces are: Torino, Valle d'Aosta, Genova, Varese, Milano, Bergamo, Brescia, Trento, Verona, Venezia, Padova, Bologna, Ancona, Firenze, Perugia, Roma, Benevento, Napoli, Salerno, l'Aquila, Bari, Taranto, Potenza, Catanzaro, Palermo and Cagliari.

Figure C.9: Cultural Intolerance Parameters - Robustness with Fertility Residuals

(a) Migrants towards Natives (Demand parameters)

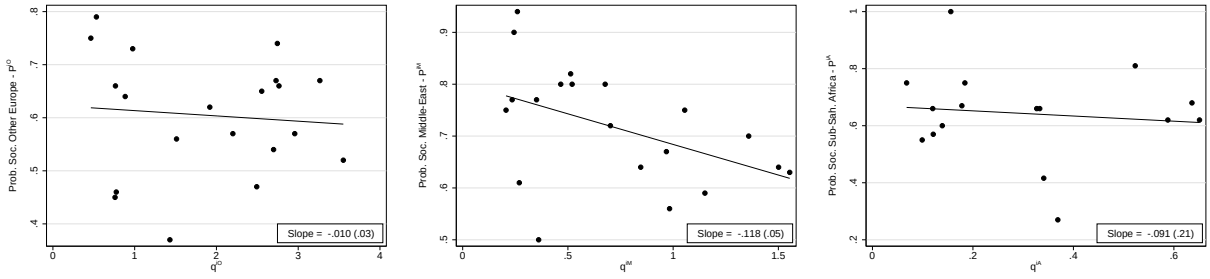
(b) Natives towards Migrants (Supply parameters)



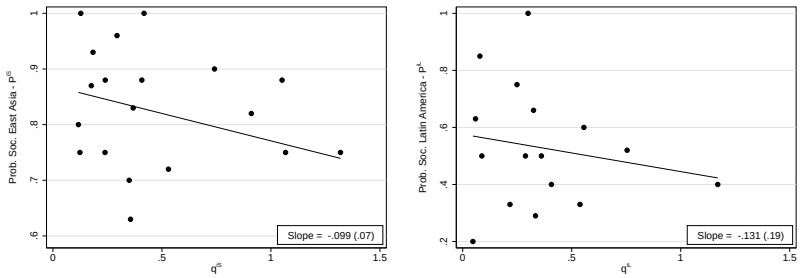
Notes: This figure reports parameter estimates of the cultural intolerance of immigrants versus natives ΔV_i^n (panel a) and natives versus immigrants ΔV_n^i (panel b) for all cultural-ethnic groups i. The blue bars report baseline estimates. The grey bars, instead, report estimates exploiting fertility residuals from a linear regression model, to control for systematic differences in observables across households, in terms of marital duration, age at marriage of spouses, as well as education and labor characteristics.

Figure C.10: Minorities Socialization Probabilities and Horizontal Socialization

(a) Other Europe (b) North Africa-Middle East (c) Sub-Saharan Africa

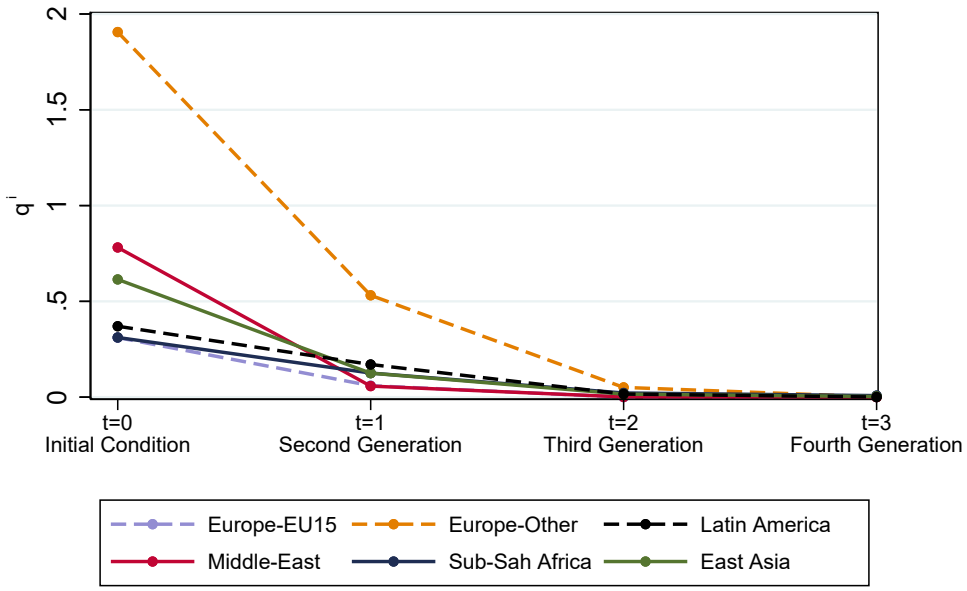


(d) East Asia (e) Latin America



Notes: This figure shows the average socialization probability of each minority group, over the correspondent population share, q^i (in percentage), for all i by region of residence (average rate over the time period). The substitution pattern displayed by Europe-EU15 minority is in line with the other minorities. However, due to sample limitations and in compliance with the ADELE Laboratory agreement, we were not allowed to export the graph.

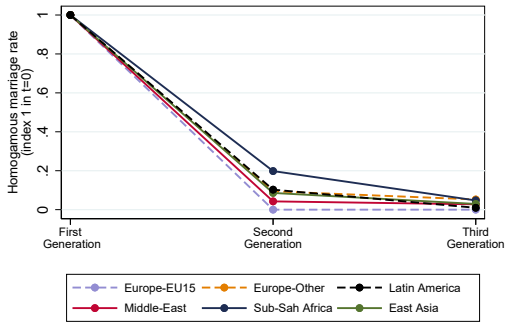
Figure C.11: Long-run Dynamics of Cultural Traits



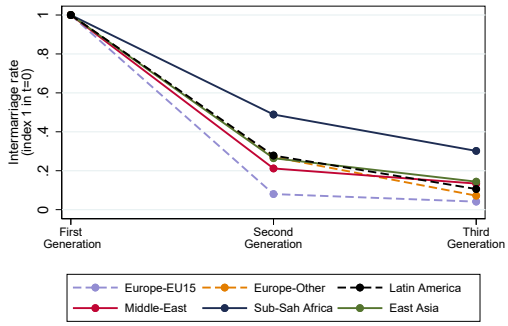
Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for all minorities, over successive generations.

Figure C.12: Dynamics of Marital Matching

(a) Homogamous



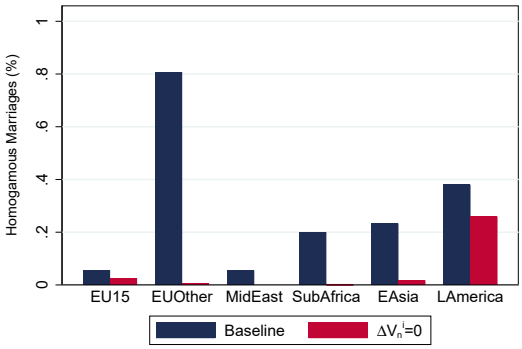
(b) Heterogamous



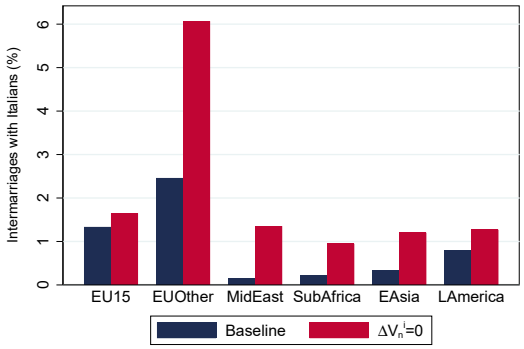
Notes: This figure shows the long-run dynamics of matching patterns for homogamous marriages (panel a) and heterogamous marriages with natives (panel b), over successive generations.

Figure C.13: Change in Matching Patterns with Italians Fully Tolerant

(a) Homogamous marriages



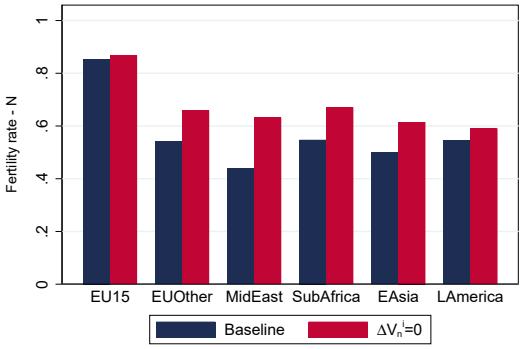
(b) Intermarriages with Natives



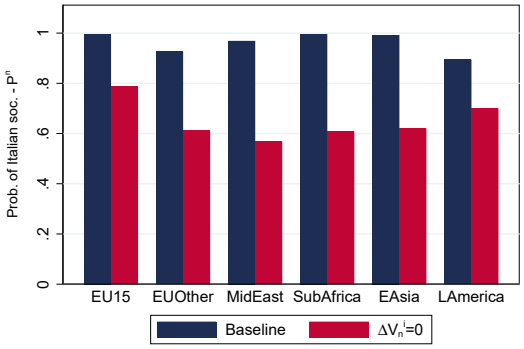
Notes: This figure shows the percentage change in homogamous (panel a) and heterogamous (panel b) marriages with full tolerance of natives towards minorities with respect to baseline, i.e., $\Delta V_n^i = 0$.

Figure C.14: Change in Intra-household Patterns with Italians Fully Tolerant

(a) Fertility

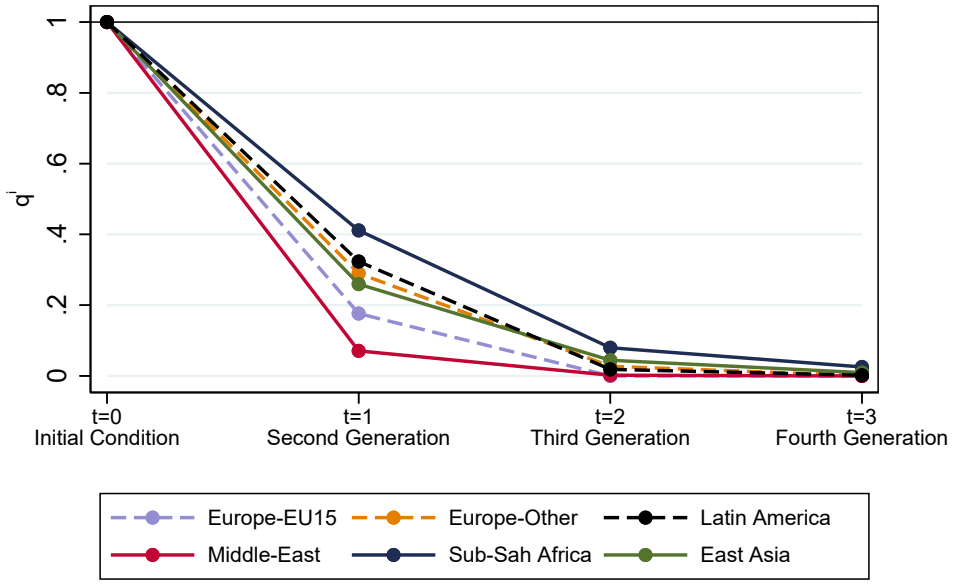


(b) Italian socialization



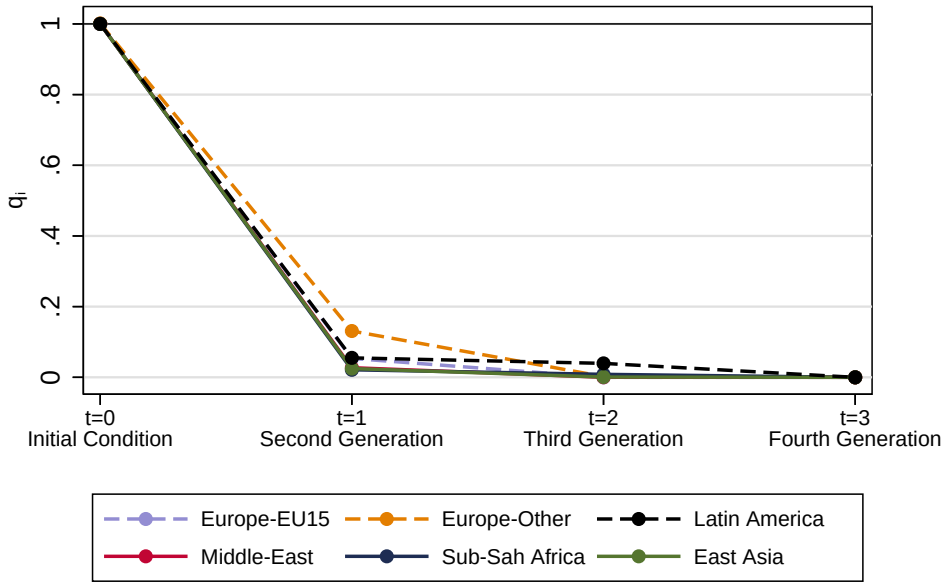
Notes: This figure shows the variation in fertility rate (panel a), Italian socialization probability (panel b) and foreign language socialization probability (panel c) in intermarriages with natives at the baseline and in case of complete tolerance of Italian majority towards minorities, i.e., $\Delta V_n^i = 0$.

Figure C.15: Dynamics of Cultural Traits with Italians Fully Intolerant, $\Delta V_n^i = 100$



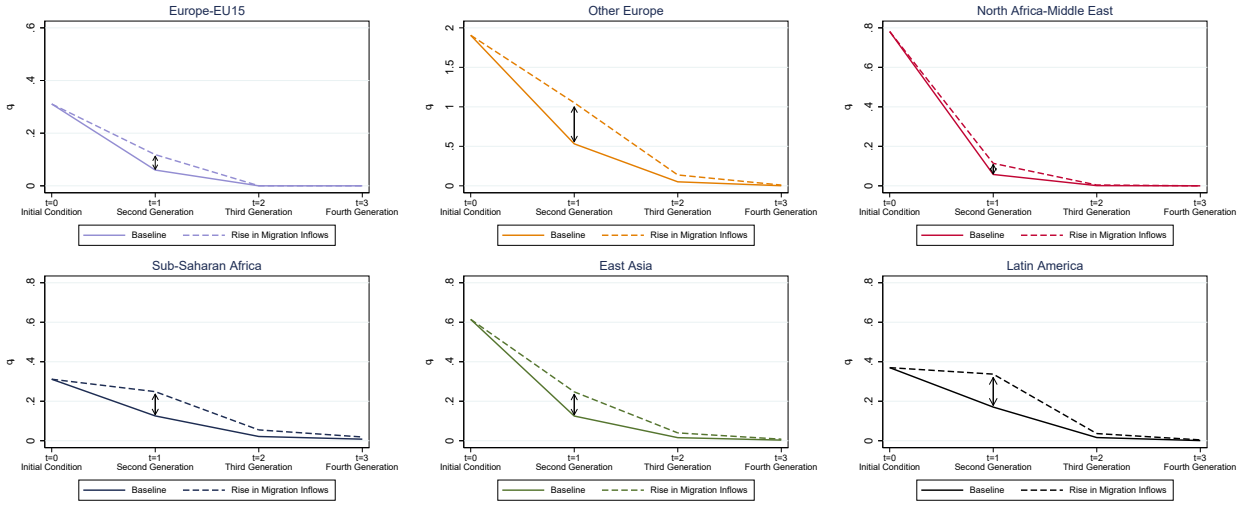
Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for all minorities i, over successive generations, assuming the case of complete intolerance of Italian majority towards minorities (q_t^i index to 1 in $t = 0$).

Figure C.16: Dynamics of Cultural Traits with Minorities Fully Tolerant, $\Delta V_i^n = 0$



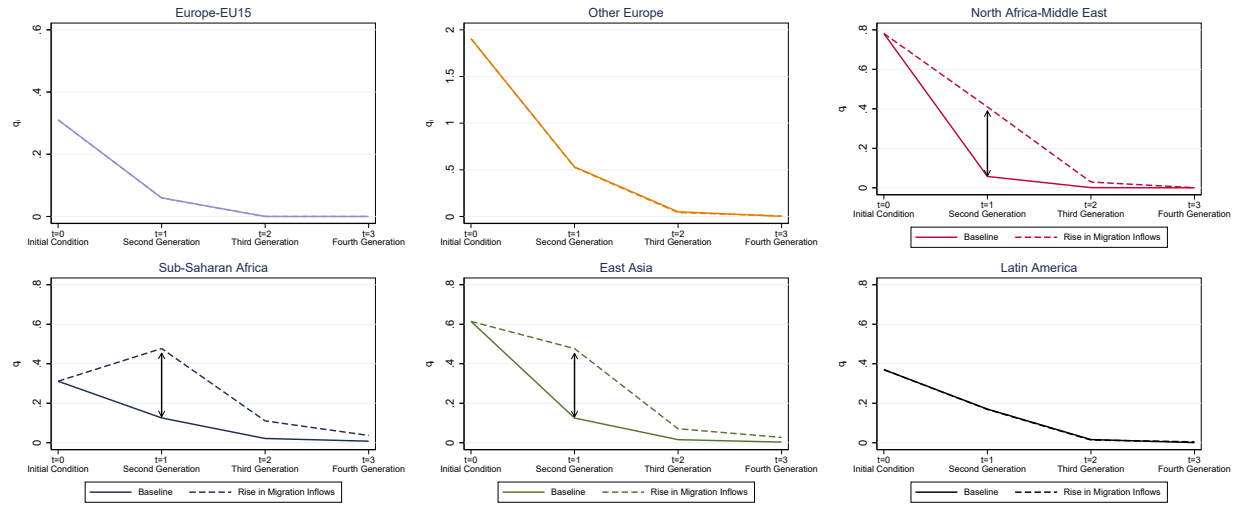
Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for all minorities i, over successive generations assuming the case of complete tolerance of minorities towards Italian culture (q_t^i index to 1 in $t = 0$).

Figure C.17: Long-run Dynamics with Proportional Raise in Migration Inflows



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations. The solid line represents the dynamics at the baseline, while the dash line represents the dynamics after doubling the share of second-generation minorities, proportionally for all minority groups. Black arrows highlight the exogenous rise in inflows for all second-generation immigrants.

Figure C.18: Long-run Dynamics with Raise in Specific Minorities Inflows



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations. The solid line represents the dynamics at the baseline, while the dash line represents the dynamics after doubling the share of second-generation North Africa-Middle East, Sub-Saharan Africa and East Asia minorities. Black arrows highlight the exogenous rise in inflows for North Africa-Middle East, Sub-Saharan Africa and East Asia second-generation immigrants.

Table C.10: Italian Language Socialization and Educational Outcomes - Robustness

	(1)	(2)	(3)	(4)	(5)
Panel A. Dep. var: Reading standardized test score, 5th grade					
Italian at Home	0.371***	0.298***	0.203***	0.202***	0.193***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	337096	336369	330739	330778	330739
R-squared	0.055	0.079	0.144	0.142	0.210
Panel B. Dep. var: Math standardized test score, 5th grade					
Italian at Home	0.239***	0.192***	0.109***	0.110***	0.103***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	352895	352121	345980	346021	345980
R-squared	0.036	0.048	0.098	0.100	0.190
Province and Cohort FE	Yes	Yes	Yes	Yes	Yes
Student Controls	No	Yes	Yes	Yes	Yes
Family Controls	No	No	Yes	Yes	Yes
School FE	No	No	No	No	Yes
Panel C. Dep. var: Choosing the high-track, 10th grade					
Italian at Home	0.064***	0.051***	0.027***	0.027***	
	(0.003)	(0.003)	(0.003)	(0.003)	
Observations	93000	92477	90656	90691	
R-squared	0.018	0.023	0.058	0.059	
Province and Cohort FE	Yes	Yes	Yes	Yes	
Student Controls	No	Yes	Yes	Yes	
Family Controls	No	No	Yes	Yes	

Notes: This table shows how our measure of Italian linguistic socialization (Italian at home) influences the achievement and educational choices of immigrant students. The dependent variables include the reading or math reading standardized test score in grade 5 in Panels A and B, respectively, and a dummy equal to one for students choosing the high-track (academic or technical schools) and zero otherwise in Panel C. The sample includes all students with at least one immigrant parent. Student controls include gender, regular schooling, a dummy for first generation immigrants, and a dummy for kindergarten. Family controls include mother's and father's education and a set of dummies for deciles of the socio-economic status distribution. Province and cohort fixed effects included in all specifications. School fixed effects included in column 5. Robust standard errors clustered at school level are in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.11: Italian Language Socialization and Additional Measures of Integration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Social network and language			Educational achievement			Italian language proficiency				
Dep. var.:	Speaking ITA w/ school mates	Having Italian friends	Speaking ITA w/ friends	High education	Pass all years	Aspiration university	Ability in Italian Reading	Writing	Speaking	Dialogue	Media
Italian at Home	0.077***	0.164***	0.249***	0.080***	0.039***	0.065***	0.141***	0.155***	0.142***	0.137***	0.139***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Observations	2,661	2,661	4,273	8,007	2,927	1,661	4,273	4,273	4,273	4,273	2,151
R-squared	0.099	0.154	0.181	0.082	0.065	0.112	0.124	0.126	0.144	0.150	0.192
Dep. var. mean	0.948	.328	0.838	0.518	0.909	0.533	.723	.712	.797	.803	.802
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows estimates of the correlation between our measure of Italian linguistic socialization (Italian at home) and various measures of socio-cultural integration concerning social networks in columns 1-3, educational achievement and aspiration in columns 4-6, and proficiency in the Italian language in columns 7-11. The sample is restricted to children and young adults (less than 25 years old), living with their parents at the time of the interview. The dependent variables include in column 1 an indicator for whether the child speaks Italian with his school mates; in column 2 an indicator for whether the child has at least some Italian friends out of the school; in column 3 an indicator for whether the child speaks Italian with his friends out of the school; in column 4 an indicator for high educational attainment (above high school); in column 5 an indicator for having passed all academic years; in column 6 an indicator for aspirations to university enrollment; in columns 7-11 a series of indicators for very good Italian proficiency in reading, writing, speaking, comprehension of interpersonal conversation and comprehension of media (television and radio newscast). Unconditional means of the dependent variables are reported below. All specifications control for province fixed effects, as well as age and gender fixed effects. Standard errors clustered at the province level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

D Identification

We provide here a proof of identification in our model. It is convenient to restrict ourselves to the set-up of the model in Section 3. In particular, we restrict to dichotomous cultural-ethnics trait, natives and immigrants (n, i). Furthermore, we assume that cost functions are quadratic, which in our set-up implies that τ and N can be solved for in closed form. The identification proof is easily extended to the more general model we estimate in Section 4. In fact, allowing for multiple cultural groups adds over-identifying restrictions to the problem. Furthermore, we do not exploit, in the following, geographical variation across $R = 20$ regions in the data, which also adds over-identifying restrictions to the problem.

We will show that the model is over-identified given the available data by proceeding recursively, solving the moment equations for the parameters to be estimated as a function of the data (given the postulated functional forms). More specifically, the data contains: the population fraction of immigrants q^i; the socialization probabilities P^c_{hj}, the divorce rates π_{hj}, the fertility rates N_{hj}, and the marriage gains G_{hj} (which are a transformation of marriage fractions).

From the socialization moments, we identify the cultural intolerances $\frac{\Delta V^{ns}_{is}}{\sigma_\tau}$, $\frac{\Delta V^{is}_{ns}}{\sigma_\tau}$. With quadratic socialization costs, $c(\tau) = \frac{1}{2}\sigma_\tau \tau^2$, when the socialization effort is $\tau = (\tau_m, \tau_f)$ and divorce is $d = 0$, the socialization moments for homogamous immigrant households with skilled parents are:

$$P^i_{isis} = 2\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i) + (1 - 2\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i))q^i.$$

This moment induces a linear equation which is independent of γ_i, S_i, and hence it is uniquely solved for $\frac{\Delta V^{ns}_{is}}{\sigma_\tau}$, taking P^i_{isis} from data.[67] The corresponding moments for $P^i_{isiu}, P^i_{iuis}, P^i_{iuiu}$ induce linear equations in γ_i, given $\frac{\Delta V^{ns}_{is}}{\sigma_\tau}$.

$$P^i_{iuis} = (1+\gamma_i)\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i) + (1 - (1+\gamma_i)\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i))q^i.$$

Same as above (because symmetric):

$$P^i_{isiu} = (1+\gamma_i)\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i) + (1 - (1+\gamma_i)\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i))q^i.$$

$$P^i_{iuiu} = 2\gamma_i\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i) + (1 - 2\gamma_i\frac{1}{\sigma_\tau}\Delta V^{ns}_{is}(1-q^i))q^i.$$

[67] The corresponding equation for $d = 1$ is linearly dependent from the first.

Therefore each of these equations is uniquely solved for γ_i, given $\frac{\Delta V_{is}^{ns}}{\sigma_\tau}$ from the previous equation and $P^i_{isiu}, P^i_{iuis}, P^i_{iuiu}$ from data.

Consider now the socialization moments for heterogamous households with $d = 0$.

$$\begin{aligned} P^i_{isns} =& \tau_m + (1 - \tau_m - \tau_f)q^i, \\ \tau_m =& \frac{1}{\sigma_\tau}(\Delta V^{ns}_{is} - \Delta V^{is}_{ns})(1 - q^i),\ \tau_f = 0 \text{ if } (\Delta V^{ns}_{is} - \Delta V^{is}_{ns}) > 0 \\ \tau_f =& \frac{1}{\sigma_\tau}(\Delta V^{is}_{ns} - \Delta V^{ns}_{is})q^i,\ \tau_m = 0 \text{ if } (\Delta V^{is}_{ns} - \Delta V^{ns}_{is}) > 0 \end{aligned} \qquad \text{(Het. in)}$$

$P^i_{isns} > q^i$ if $\tau^m > 0$ (which implies $\tau^f = 0$). Hence $\frac{\Delta V^{ns}_{is} - \Delta V^{is}_{ns}}{\sigma_\tau} > 0$. Once this sign is identified, then, the moment for P^i_{nsis} is linear in $\frac{\Delta V^{ns}_{is} - \Delta V^{is}_{ns}}{\sigma_\tau}$. Since we previously identified $\frac{\Delta V^{ns}_{is}}{\sigma_\tau}$, this identifies $\frac{\Delta V^{is}_{ns}}{\sigma_\tau}$.

The moment for P^i_{nuis} is as follows.

$$\begin{aligned} P^i_{nuis} =& \tau_f + (1 - \tau_m - \tau_f)q^i, \\ \tau_m =& \frac{1}{\sigma_\tau}(\gamma_n \Delta V^{is}_{ns} - \Delta V^{ns}_{is})(1 - q^i),\ \tau_f = 0 \text{ if } (\gamma_n \Delta V^{is}_{ns} - \Delta V^{ns}_{is}) > 0 \\ \tau_f =& \frac{1}{\sigma_\tau}(\Delta V^{ns}_{is} - \gamma_n \Delta V^{is}_{ns})q^i,\ \tau_m = 0 \text{ if } (\Delta V^{ns}_{is} - \gamma_n \Delta V^{is}_{ns}) > 0 \end{aligned} \qquad \text{(Het. ni)}$$

It is linear in $\frac{\gamma_n \Delta V^{ns}_{is} - \Delta V^{is}_{ns}}{\sigma_\tau}$ and hence identifies it. Since we previously identified $\frac{\Delta V^{ns}_{is}}{\sigma_\tau}$, and $\frac{\Delta V^{is}_{ns}}{\sigma_\tau}$, this identifies γ_n.[68]

As a consequence, we have identified $\frac{\Delta V^{ns}_{is}}{\sigma_\tau}$ and $\frac{\Delta V^{is}_{ns}}{\sigma_\tau}$. Note that this is the case, independently, for every region r, as long as the distribution of the population q^i varies across region. Assuming that $\Delta V^{ns}_{is}, \Delta V^{is}_{ns}, \sigma_\tau$ are constant across region r, we cannot exploit the variation across regions to identify them separately.

From the divorce moments, we identify the means of the distribution of the marriage quality shock θ, a_{hj}; the utility of a child within marriage δ, the socialization costs σ_τ, and the values of education S_i, S_n. Recall that δ denotes the utility of having a child within marriage; while θ denotes the marriage quality shock. The distribution of θ_{hj} has a mean that depends on the household type hj. θ_{hj} follows a generalized logistic distribution cumulative distribution function $F(\theta_{hj})$, with location a_{hj} and scale parameter b. We set $b = 1$. The parameters a_{hj} are identified independently from any other parameter

[68] In fact, γ_n is over-identified as it can also be identified by the moment for P^i_{nuiu}, which is linear in $\frac{\gamma_n \Delta V^{ns}_{is} - \gamma_i \Delta V^{is}_{ns}}{\sigma_\tau}$.

in the model with data on the dissolution probability of couples *without children*, $\hat{\pi}_{hj}(0)$, for all h, j; that is, a_{hj} is the unique solution to

$$F(0; a_{hj}, 1) = \hat{\pi}_{hj}(0), \text{ for all } h, j.$$

The moments for the probability of divorce, for any household type hj, are:

$$\pi_{hj}(N_{hj}) = F(N_{hj}u_{hj}(\tau(1), 1) - N_{hj}u_{hj}(\tau(0), 0) - N_{hj}\delta). \tag{Div}$$

Consider first equation Div for $h = ns;\ j = ns$. Normalizing $V = 100$,

$$\begin{aligned} u_{nsns}(\tau, 0) =& 2V - 2P^u_{nsns}S_n \\ u_{nsns}(\tau, 1) =& 2V - 2P^u_{nsns}S_n, \end{aligned}$$

and hence the equation Div is independent of S_i, S_n. It is only a function of δ, which is then identified. Consider now equations Div for $h = is;\ j = is$ and for $h = iu;\ j = is$, where

$$\begin{aligned} u_{isis}(\tau, 0) =& 2VP^i_{ii}(\tau, 0) + 2V^n_i(1 - P^i_{ii}(\tau, 0)) - 2P^u_{isis}S_i - \frac{1}{2}\sigma_\tau(\tau_m + \tau_f)^2 \\ u_{isis}(\tau_f, 1) =& 2VP^i_{ii}(\tau_f, 1) + 2V^n_i(1 - P^i_{ii}(\tau_f, 1)) - 2P^u_{isis}S_i - \frac{1}{2}\sigma_\tau(\tau_f)^2, \\ u_{iuis}(\tau, 0) =& (1 + \gamma_i)VP^i_{ii}(\tau, 0) + 2V^n_i(1 - P^i_{ii}(\tau, 0)) + (1 - \gamma_i)S_i - 2P^u_{iuis}S_i - \frac{1}{2}\sigma_\tau(\tau_m + \tau_f)^2 \\ u_{iuis}(\tau_f, 1) =& (1 + \gamma_i)VP^i_{ii}(\tau_f, 1) + 2V^n_i(1 - P^i_{ii}(\tau_f, 1)) + (1 - \gamma_i)S_i - 2P^u_{iuis}S_i - \frac{1}{2}\sigma_\tau(\tau_f)^2. \end{aligned}$$

These equations are a function of σ_τ, $\frac{\Delta V^{ne}_{ie}}{\sigma_\tau}$, δ, S_i and hence identify σ_τ and S_i. Notice that, since we previously identified $\frac{\Delta V^{ne'}_{ie}}{\sigma_\tau}$, $\frac{V^{ie'}_{ne}}{\sigma_\tau}$, for any e, e', we have now identified the cultural intolerances $\Delta V^{ne'}_{ie}, \Delta V^{ie'}_{ne}$.

From the fertility moments, we identify the fertility cost σ_N. Assuming quadratic fertility costs, $\kappa(N) = \frac{1}{2}\sigma_N N^2$, the fertility moments for any household type hj, are:

$$N_{hj} = \frac{1}{\sigma_N}\left(\pi_{hj}(N_{hj})u_{hj}(\tau(1), 1) + (1 - \pi_{hj}(N_{hj}))(\delta + u_{hj}(\tau(0), 0))\right).$$

Consider these moments for $h = ns; j = ns$ and for for $h = is; j = is$. Using these equations uniquely identify σ_N and S_n with data on N_{nsns}, N_{isis}, given the parameters previously identified. The moments for heterogamous households allow for over-identification.

From the marriage fractions and marriage gains, we identify the outside options of marriages, $\omega_{h_z}, \omega_{j_z}$ for $z \in hom, het$. For each household type hj, the moment equations

for the implied marital surplus net of outside options are

$$U_{hj} - \omega_{h_z} - \omega j_z = G_{hj}$$

where gains from marriage are defined as:

$$G_{hj} = \log \frac{(\mu_{hj})^2}{\mu_{h.} \cdot \mu_{.j}}; \tag{9}$$

and U_{hj} is the utility of household hj, which is identified given the parameters of the model we have already identified. In turn, the marriage distribution μ_{hj} and fractions of singles $\mu_{h.}, \mu_{.j}$ are data. These moments then depend on the parameters already identified and linearly on ω_{c_z}, which they identify. In this particular case, with traits and two family types, we have 2×2 outside options to be estimate, and 4 gains. So ω_{c_z} are identified.