ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 73/25

Marriage, Fertility, and Cultural Integration in Italy

Alberto Bisin, Giulia Tura

www.rfberlin.com

SEPTEMBER 2025

Marriage, Fertility, and Cultural Integration in Italy

Authors

Alberto Bisin, Giulia Tura

Reference

JEL Codes: D1, J12, J13, J15

Keywords: Marital Matching, Fertility, Cultural Transmission, Integration

Recommended Citation: Alberto Bisin, Giulia Tura (2025): Marriage, Fertility, and Cultural Integration in Italy. RFBerlin Discussion Paper No. 73/25

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Marriage, Fertility, and Cultural Integration in Italy*

Alberto Bisin† Giulia Tura‡

September 10, 2025

Abstract

We study cultural integration as an equilibrium outcome of marital matching along cultural and education lines and intra-household investment decisions regarding fertility and cultural socialization. We show that our marriage model allows us to identify cultural-ethnic group specific investment parameters as well as spousal preferences for marital matching. Structural estimates fit the data well and reveal a strong demand to preserve cultural identity on the part of immigrants as well as limited acceptance of the immigrants' cultural diversity on the part of natives. Furthermore, these estimates reveal a substantial heterogeneity of the parental value of children's education across cultures. Nonetheless, our estimates imply substantial - though heterogeneous - cultural integration rates across immigrant groups in simulations.

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

*We thank Jérôme Adda, Pierre-André Chiappori, Alex Citanna, Alfred Galichon, Sanjeev Goyal, Yujung Hwang, Ilse Lindenlaub, Corinne Low, Costas Meghir, Giovanni Peri, 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 and Finance, LUISS University, 00198 Rome, Italy. email: gtura@luiss.it.

1 Introduction

The recent surge in migration flows into Western countries represents one of the most contentious political and socio-economic phenomena of the last decades. Widespread restrictive immigration policies are motivated in part by the perceived externalities immigration imposes on natives and by the perceived sluggishness of the cultural integration process of immigrants.[1]

To better understand the political economy of immigration, we study the dynamics of cultural integration as the outcome of a process of marriage formation and intra-household investment decisions over fertility and cultural socialization. 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.

We introduce a structural model of the marriage market in which spouses i) match along the cultural-ethnic dimension as well as along the educational dimension and ii) solve a collective household investment decision problem determining fertility and children's socialization. Furthermore, iii) spouses choose endogenously whether to divorce and this decision affects children's socialization. More specifically, parents care about their children's education and the value of education depends on the cultural-ethnic group of the parent. Parents also care about their children's cultural identity and these preferences vary with the parents' cultural-ethnic group and their own education. For simplicity, we assume that the transmission of education is exogenous and depends on parental education and cultural-ethnic group. On the other hand, we model cultural socialization as the outcome of parental choice. Parents who value their cultural identity care about socializing their children and are endowed with socialization technologies to transmit their own cultural-ethnic traits to children. Thus, parents choose to exert a direct socialization effort to affect their children's process of cultural identity formation. But effort requires costly parental resources, for example, time spent with children, private school tuition, selection of residential neighborhoods, ethnic networks,

[1] Negative labor market effects of immigration on natives are far from well-documented; see Card (1990), Borjas (2003), Bisin and Zanella (2017), Dustmann et al. (2017), and Caiumi and Peri (2024). For empirical evidence on the speed of cultural integration see e.g., Algan et al. (2012) and Abramitzky and Boustan (2017).

and so on. Furthermore, socialization technologies vary between culturally homogamous and heterogamous marriages. For instance, homogamous households where parents share the same cultural traits enjoy a more efficient socialization technology in transmitting their shared traits than families where parents do not share the same culture. Similarly, fertility and separation choices depend on the education and 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 separation patterns, across households' education and cultural-ethnic characteristics, which can be taken to the data.

To this end, we exploit rich administrative marriage data providing information on the cultural-ethnic identity of immigrants in Italy.[2] 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 as well as the universe of births and separations registered in the same period. In addition, we recover a measure of parental cultural socialization of children from the *Condition and Social Integration of Foreign Nationals* Survey (2011-2012). In a reduced form analysis, we show that the systematic dependence of marriage, fertility, socialization, and separation patterns induced by our model is in line with the stylized facts and descriptive patterns in the data. We document the relevance of cultural-ethnic attributes, as central determinants of the observed marriage, fertility, education, and socialization patterns. We observe strong positive assortative matching along cultural-ethnic lines, even stronger than along education lines. Moreover, we uncover sizable differences between culturally homogamous and heterogamous marriages in terms of fertility and socialization, with homogamous marriages displaying systematically higher childbirth and socialization investments. Finally, we show that the effects of parental education on fertility and socialization are systematically mediated by the cultural sorting of the marriage. This reduced form analysis, however, does not allow us to identify the determinants of the observed process of cultural integration of immigrants and is silent about any underlying possible interactions between individuals' preferences along cultural and educational lines. We thus estimate the parameters of our structural model.

Our structural analysis contributes to the literature on both methodological and empirical dimensions. Methodologically, we show that our marriage model, with endogenous parental investments and endogenous divorce choices, can be transparently identified by exploiting cross-sectional variation in outcomes across household types by cultural groups and education levels. We formally demonstrate that it is possible to identify cultural-ethnic group–specific

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

investment parameters and spousal preferences for marital matching. By embedding parental investments into the marriage matching framework, our model induces a parametrization of individual spouses' preferences that is identified using data on marriage patterns, fertility, divorce, and cultural socialization. This represents a significant advance over the existing marriage matching literature, pioneered by Choo and Siow (2006b). In that literature, match-specific surplus parameters (i.e., the joint gains from marriage) are identified from observed matching patterns, but the individual preferences of each spouse cannot be identified without observing transfers across spouses.

Empirically, we show that our structural model fits the data well. We estimate a rich set of parameters including i) *cultural intolerance parameters* - a measure of the parental preference for the inter-generational transmission of the culture of a specific cultural-ethnic group;[3] ii) the *value of children's education* - by cultural-ethnic group - and the relative strength of the cultural preferences of parents with high and low education - also by cultural-ethnic group; and iii) preferences for fertility, fertility and cultural socialization costs, the outside options to marriage, and the relative effects of marriage on spouses' economic opportunities.

Estimated cultural intolerance parameters are positive, asymmetric, and highly heterogeneous across cultural-ethnic groups. In other words, preferences for cultural identity dominate economic incentives for 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 that of Europeans. On the other hand, we also estimate high cultural intolerances on the part of 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 the value of education is highly heterogeneous by cultural-ethnic groups and parental cultural intolerances tend to be higher for immigrant parents with low education.

By simulating our model at the estimated parameters, we can investigate the inter-generational evolution of the distribution of the population by cultural traits. Despite high cultural intolerance estimates, all cultural-ethnic minorities are simulated to integrate into the Italian majority. The cultural 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 into the native Italian culture, is 75% in one single generation. However, the pace of convergence is heterogeneous across cultural-ethnic groups, with slower integration rates

[3]The cultural intolerance of immigrants identifies the demand side of integration at equilibrium, while the cultural intolerance of natives identifies the supply of cultural acceptance towards immigrants' cultural diversity in the society.

characterizing in particular immigrants from East Asia, Sub-Saharan African minorities, and especially Latin America (which reaches full convergence only after four generations). Interestingly, the patterns of cultural integration across ethnic groups in the simulations are not only the result of the estimated cultural intolerances but also of substantial heterogeneity in cultural homogamy and fertility rates. The inter-generational dynamics of immigrants' education levels - an indication of their economic integration - essentially display constant immigrants' education levels over successive generations.

To examine in more depth the mechanisms driving cultural integration at equilibrium, we perform a counterfactual exercise. More specifically, we study how integration responds to variations i) in the supply of acceptance of the immigrants' cultural diversity on the part of natives as well as ii) in the demand on the part of immigrants to preserve their cultural identity. First, setting the cultural intolerance of natives with respect to all ethnic minorities equal to zero, we observe a significantly lower integration rate of immigrants. When natives are fully tolerant, in fact, immigrants in heterogamous marriages with natives can achieve higher socialization rates. At equilibrium, heterogamy increases and so does socialization and fertility in intermarriages, reducing integration overall. Second, strengthening the dominance of cultural identity in the demand of immigrants, which could for instance be due to a reduction in their economic incentives to integration, we find that cultural convergence is substantially accelerated. This result is the outcome of lower participation in the marriage market and lower fertility among immigrants, motivated by the fact that a stronger attachment to their cultural identity (and weaker economic incentives to integrate) makes marriage costlier in terms of socialization effort.

1.1 Related literature

Our paper combines insights from the literature on the inter-generational transmission of cultural traits with those of family economics on the estimation of (marital) matching preferences. Our methodological contribution is twofold. On the one side, 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).[4] In our model, marital utilities emerge endogenously as a function of fertility and intra-household inter-generational socialization choices with respect to culture - along the

[4]See Choo and Siow (2006b); Chiappori et al. (2009); Dupuy and Galichon (2014); Choo (2015); Ahn (2024); 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.

lines of Bisin et al. (2004) - and also with respect to education.[5] We account then 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 cultural-ethnic traits and educational preferences crucial dimensions of marital matching (Bisin et al., 2004; Ciscato and Weber, 2020; Chiappori et al., 2018). On the other side, the paper advances the literature on marital matching, pioneered by Choo and Siow (2006b), who first provided exact identification of marital preferences, that is, marriage gains at the household level. By introducing endogenous parental investments in fertility, cultural socialization, and separation choices within a matching model, we prove that we can separately identify the preferences of individual spouses, which jointly shape the marriage equilibrium and household decisions.

In terms of empirical research question, this paper fits into the large literature on the cultural integration of immigrants. Several of these studies concentrate on the immigrants' demand to preserve their cultural identity, by exploring socialization e.g., 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, 2020); 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 Peri, 2016; Dustmann et al., 2016), and their consequent political reactions (Dustmann et al., 2019; Tabellini, 2020; Campo et al., 2023). With respect to this empirical literature on cultural integration, our main contribution relies on the ability to identify and estimate separately the demand of immigrants to preserve their cultural identity as well as the supply of acceptance of the immigrants' cultural diversity on the part of natives. Furthermore, our approach allows us to identify and estimate a parametrization of the fundamental interactions between the cultural traits and the education of the parents in their preferences over spouses and 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.

[5]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.

2 Marriage by cultural group and education

This section introduces our empirical analysis of marriage by cultural-ethnic group and spouses' education, along with data sources. We document the main marriage patterns and we relate them to relevant outcomes such as fertility, separation, children's cultural socialization and education. While this analysis shows the central role of cultural and educational attributes in the marriage market, it is necessarily silent on the identification of the fundamental mechanisms underlying immigrants' integration outcomes. We therefore interpret this section as motivation for the structural analysis of marriage that follows.

2.1 Data

We use administrative individual-level data from the Italian Statistical Institute (ISTAT, ADELE Laboratory), covering the universe of marriages celebrated in Italy from 1995 to 2012. Crucially, marriage data account only for legal marriages and only when celebrated in Italy. We complement marriage data with the universe of births (from municipality birth registries) and legal separations (from civil court chancelleries) registered during the same period. Our final sample consists of over 4 million marriages – 94% of all marriages during the period. The fertility rate among these marriages is 69.6%, with an average of 1.54 children per household, while 7% of marriages end in separation within the first years.

Marriage records include detailed socio-demographic information on both spouses, such as date and age at marriage, education level, employment status, municipality of residence, and for immigrants, nationality and country of origin. A comprehensive description of the data sources and variables of interest is provided in Appendix A. For cultural and ethnic classification, we distinguish between Italians, the native majority group, and six immigrant minority groups, aggregated by country of origin: European (EU15 countries), Other European, North African and Middle Eastern, Sub-Saharan African, East Asian, and Latin American; see details in Figure C.1. In terms of education, we group individuals into two categories: high education (at least some high school) and low education (no high school). This binary classification captures the most relevant variation in educational attainment in our data; Figure C.2 illustrates the distribution of educational attainment.[6] Geographically, we divide Italy into 20 distinct local marriage markets, corresponding to the administrative regions in the country. In addition, we use individual-level Italian Census data from 2001 and

[6]In a comparative perspective, Italy ranks among the lowest in Europe for tertiary education attainment. Despite recent increases in public investment, only about 20% of Italians had completed at least a bachelor's degree in 2023. Meanwhile, 42% had completed upper secondary or post-secondary non-tertiary education, and 44% had attained only primary or lower secondary education (Eurostat data).

2011 to estimate the distribution of unmatched adult men and women, by cultural-ethnic group, education, and region.

Finally, we proxy cultural-ethnic transmission through language, using "speaking Italian at home" as an indicator of unsuccessful transmission, since ethnic identity and language are closely related culturally transmitted traits (Dustmann, 1997; Casey and Dustmann, 2008; Ginsburgh and Weber, 2011; Clots-Figueras and Masella, 2013; Fouka, 2020). Indeed, speaking Italian at home is consistently associated with weaker ethnic identity and stronger attitudes toward social integration of immigrant students; see Section 5.3 for further evidence. Data on "language spoken at home" come from the *Condition and Social Integration of Foreign Nationals Survey*, conducted in 2011–2012 across all Italian regions on a sample of 9,600 foreign resident families. The survey aims to provide a comprehensive picture of the socio-cultural and economic integration of foreign residents, gathering information on living conditions, behaviors, attitudes, and opinions. For our analysis, we focus on children and young adults under the age of 25 living with their parents at the time of the interview. This yields a final sample of 8,007 individuals from approximately 5,000 families, of which 86.7% have married parents, while the remainder are separated or divorced. We classify individuals as "speaking Italian at home" if they report speaking Italian within their family; otherwise, we assign them the mother tongue learned in early childhood.

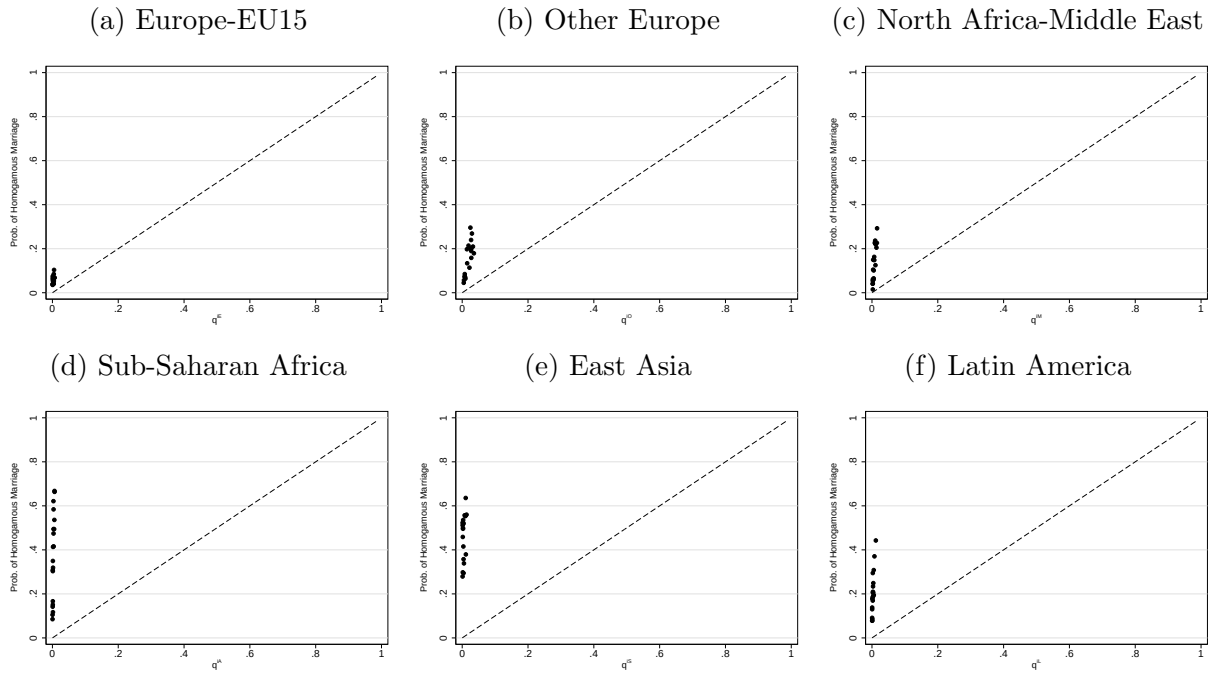
2.2 Marriage

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%. 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, which are represented by the 45-degree line.

Beyond cultural origins, individuals also match on 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 reflects this pattern; in 77 percent of marriages in our sample, both spouses have the same education level (46 percent high and 31 low). See Figure C.3 for a heatmap of the distribution of marriages (in log) by ethnic group and spouses' education.

The marriage rates in our data suggest that both cultural and educational factors constitute an important determinant of marriage choices. However, these rates depend on the

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 q^i (reported on the horizontal axis). The distribution of q^i is averaged over the period 1995-2012.

distribution of men and women by cultural group and education in the marriage market under consideration. To control for the uneven distribution of cultural traits and education in the population and for potential gender imbalances, we turn to the standard measure in the literature, *gains from marriage*. 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 from population Census data. Besides controlling for the distribution of men and women by their relevant traits, gains from marriage can be interpreted as a measure of the utility gains associated with marriage with respect to the outside options of remaining single, for each marriage type (Choo and Siow, 2006b).[7] Figure C.3 also depicts the distribution of gains from marriage. As suggested by marriage rates, we observe that higher gains are concentrated along the main diagonal, corresponding to homogamous marriages, indicating strong positive sorting along both cultural and educational lines.

As a first attempt at disentangling cultural and educational factors in marriage, Table

[7]Since Choo and Siow (2006b), gains from marriage have been studied in different contexts of interest; see references in Footnote 4. Gains from marriage are defined with respect to a predetermined set of marriage types; in our case with respect to cultural and educational characteristics of the spouses. We refer to Section 3.1 for a formal discussion of gains from marriage and their identification from marriage data.

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 and low education (columns 3 and 4).[8] Finally,

Table 1: Gains from Marriage by Culture and Education

	(1)	(2)	(3)	(4)	(5)
Dep. var:	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	2,456	2,456	2,456	2,456	2456
N. Marriages	4,231,283	4,231,283	4,231,283	4,231,283	4,231,283
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: This Table shows the results of the regression of gains from marriage - computed as in equation (2), across cells defined by region, spouses' cultural ethnic-group and education - on various (combinations of) explanatory variables. In columns (1) and (2), the explanatory variables are dummies for homogamous marriages by culture (*Homogamous by culture*) and education (*Homogamous by education*). Columns (3)-(5) include dummies for homogamous marriages by education, distinguishing marriages with high (*High-high educ*) and with low (*Low-low educ*) education of both spouses. Column (5) includes interaction variables. All specifications include marriage market (at the regional level) fixed effects, as well as wife educational level and cultural-ethnic group of origin fixed effects. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.
Source: ISTAT, marriage data (1995-2012) and individual Census data.

[8] In Table C.2, we report on a further empirical exercise which strengthens this result by showing 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) - even after conditioning on education. This correlation is statistically significant, consistently for different measures of cultural distance.

Table 1 documents a positive and statistically significant interaction between cultural and educational homogamy (column 5). Interestingly, homogamous cultural marriages of spouses with low education exhibit higher gains from marriage compared to similar homogamous marriages of highly-educated spouses, suggesting stronger complementarities in marriage.

2.3 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 with the observed marriage pattern - and hence with marriage gains - in the data. 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 homogamy - by culture and education - and fertility, separation rate, as a measure of lack of success in marriage, and various 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, shaped 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.

Focusing on fertility, Table 2 documents that culturally homogamous marriages display systematically higher fertility, independently of the spouses' education; see also Table C.2 panel b). The effect of education on fertility, instead, is not statistically significant on average, but it is mediated by cultural homogamy, i.e., it is positive for culturally heterogamous marriages and negative for culturally homogamous marriages; see column (5). 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. This differential role of education on fertility - mediated by the cultural matching of the marriage - appears to be a systematic feature of the data, see Figure C.4 for a graphical representation. Indeed, this is the case also for the separation rate. Table C.4 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 classifi-

Table 2: Fertility Rate by Culture and Education

Dep. var:	(1)	(2)	(3)	(4)	(5)
	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
Ethnic and Education (wife) FE	Yes	Yes	Yes	Yes	Yes

Notes: This Table shows the results of the regression of the fertility rate - in marriages defined by spouses' cultural ethnic-group and education, and by region - on various (combinations of) explanatory variables. In columns (1)-(3), the explanatory variables are dummies for homogamous households by culture (*Homogamous by culture*) and by education (*Homogamous by education*). In column (4), we include a dummy for marriages with at least one high educated spouse (*High educ, at least one*) and also its interaction with the dummy for cultural homogamous marriages (*Hom by culture* × *High educ, at least one*) in column (5). 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 the wife's high education. Results are robust to different classifications of sorting by education; see Table C.3. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.
Source: ISTAT, marriage data (1995-2012) and birth data (1995-2012).

cation for the children as for the parents and we take "speaking Italian at home" as a proxy for unsuccessful cultural socialization. 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").[9] The role of parental education is once again mediated by cultural homogamy: high parental education has a significant negative role on socialization only in culturally homogamous marriages. As for children's education, we find that while the probability that a child is highly-educated is positively correlated with the education of the parents, this correlation is stronger in culturally heterogamous rather than in culturally homogamous marriages; see Table C.6.

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

[9]Immigrant parents in homogamous marriages are, on average, half as likely to speak Italian with their child as those in heterogamous marriages with another immigrant; see Figure C.5. This aligns with other evidence using alternative measures of socialization (Dohmen et al., 2012; Fouka, 2020, among others).

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

Dep. var:	(1)	(2)	(3)	(4)	(5)
	"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
Ethnic and Education (wife) FE	Yes	Yes	Yes	Yes	Yes

Notes: This Table shows the results of the regression of the fraction of "Italian spoken at home" - in marriages defined by spouses' cultural ethnic-group and education, and by region - on various (combinations of) explanatory variables. In columns (1)-(3), the explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and by education (*Homogamous by education*). In column (4), we include a dummy for marriages with at least one educated spouse (*High educ, at least one*) and also its interaction with the dummy for cultural homogamous marriages (*Hom by culture* × *High educ, at least one*) in column (5). All specifications include marriage market (at the regional level) fixed effects, spouses' cultural-ethnic group of origin fixed effects, and a dummy for wife's high education. Results are robust to different classifications of sorting by education; see Table C.5. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.
Source: ISTAT, Condition and Social Integration of Foreign Nationals Survey (2011–2012).

with various important outcomes of marriages: their fertility and stability as well as various dimensions of children's quality.[10] Finally, the evidence indicates that the effect of parental education on various marriage outcomes is systematically mediated by the cultural sorting of the marriage. This analysis suggests interesting possible interactions between preferences along cultural and educational lines. In particular, we have documented some possibly systematic relationships between parents' culture and education. We also have found evidence regarding parental representations of the quantity-quality trade-off in fertility.

The analysis of marriage gains we pursued in this section however cannot identify the fundamental aspects of individual preferences as they are manifested in the marriage market. Notably, it cannot separate preferences ranking spouses by education level from preferences

[10]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); in economics, see Bisin and Verdier (2000); Bisin et al. (2004). On the role of children's education, see Becker (1973, 1974); Chiappori et al. (2017, 2018).

valuing educational homogamy *per se* in marriage; e.g., it cannot reveal whether highly educated individuals (in some cultural-ethnic groups) value cultural identity less, in a quality-quantity trade-off logic. Most importantly, this analysis cannot allow for a decomposition of how much marriage gains are determined by the cultural preferences of immigrants to preserve their 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, see also Vieira et al. (2022) and Fouka (2024) for comments along these lines.

Motivated by these limitations in the descriptive analysis of marriage gains, we turn to a structural model in which marriage allocations reveal the deep preferences of individuals with respect to several a priori relevant marriage outcomes - 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 a direct (exogenous) link between marriage type and children's education, calibrating it from the data. Estimating this equilibrium model of marriage, fertility, and socialization allows us to identify the distribution of cultural preferences along the cultural-ethnic dimension as well as any structural relationship between education and cultural preferences. 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.

3 A model of marriage and household 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. For simplicity, we present the theoretical model for dichotomous cultural traits c - with $c = n$ for natives and $c = i$ for immigrants - and dichotomous educational levels e - with $e = s$ for high education and $e = u$ for low education. The set of characteristics is then $C = \{ns, nu, is, iu\}$.[11]

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.[12] Households' heterogeneity builds on the individuals'. The notation hj denotes a household type where the male has characteristic h and the female j, with $h, j \in C$; $h0$

[11] In the empirical analysis, we shall allow for multiple cultural-ethnic traits of immigrants.

[12] 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.

denotes the household type composed of a single male with characteristic h and $0j$ the one composed of a single female with characteristic j. T denotes 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 outcomes of future household choices and ii) an idiosyncratic component, capturing residual idiosyncratic returns from marriage. Let m index individual males and f individual females in the population. Let ϵ_{mj} and η_{hf} denote the individual idiosyncratic preference shocks, respectively, of an individual man m for a woman of type j and of an individual woman f for a man of type h. In turn, ϵ_{m0} and η_{0f} denote the idiosyncratic component of utility associated with staying single.

Individuals observe their idiosyncratic shocks before marriage and then match along cultural-ethnic traits and educational levels, anticipating their marital utility. The total marital utility of a household of type hj - between man m with identity h and female f of identity j - is $U_{hj} + \epsilon_{mj} + \eta_{hf}$; that is, utility is additive and separable in the shocks. This separability assumption rules out any complementarity between unobserved spouses' characteristics. The total utility of a single male (resp. female) is $U_{h0} + \epsilon_{m0}$ (resp. $U_{0j} + \eta_{0f}$). Following Choo and Siow (2006b), we assume that the individual idiosyncratic shocks $\epsilon_{mj}, \eta_{hf}, \epsilon_{m0}, \eta_{0f}$ are independent and identically distributed random variable vectors with a type I extreme-value distribution (Gumbel).[13]

3.1 Matching equilibrium

Given the systematic component of marital utility U_t for all different potential matches $t \in T$, let μ_t denote the fraction of marriages of type t formed in the population, where μ_{hj} denotes the fraction of marriages composes of a male of type h and a female spouse of type j, and μ_{h0} and μ_{0j} denote the fraction of single men of type h and single women of type j, respectively. Let m_h and f_j denote males and females with trait h and j, in turn, in the marriage market.

The equilibrium notion we adopt in the marriage market requires $\{\mu_t\}_{t \in T}$ to represent (a configuration of) *stable matches* (Shapley and Shubik, 1971); that is, it requires that, given

[13] While imposing strong restrictions on agents' unobserved heterogeneity, this assumption is standard in empirical practice following Choo and Siow (2006b)'s seminal contribution and is motivated by computational simplicity. In our matching model with collective decision problems, it is crucial for tractability. We acknowledge several attempts in the literature to relax the parametric distributional restrictions on agents' heterogeneity in simpler settings, either by modifying the logit assumption or introducing gender-specific heteroskedasticity; see, for example, Dupuy and Galichon (2014); Galichon and Salanié (2022); Chiappori et al. (2017, 2019). For an alternative perspective within a fully non-parametric framework, see Gualdani and Sinha (2023).

an equilibrium assignment of matches $\{\mu_t\}_{t\in T}$, no agent would prefer to deviate from the assignment (no blocking pairs exist). It is shown in the literature that an equilibrium (stable) configuration $\{\mu_t\}_{t\in T}$ corresponds to the solution of the following constrained maximization problem (e.g., Chiappori and Salanié, 2016; Galichon, 2018):

$$\begin{aligned} &\max_{(\mu_t\geq 0)_{t\in T}} \quad \sum_{t\in T}\mu_t U_t - \varepsilon(\mu) \\ &s.t. \\ &\sum_j \mu_{hj} + \mu_{h0} = m_h \quad \forall h \in C, \\ &\sum_h \mu_{hj} + \mu_{0j} = f_j \quad \forall j \in C \end{aligned} \tag{1}$$

where, i) under our separability and distributional assumptions on the individual unobserved heterogeneity components, the expression $\varepsilon(\mu)$ 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, 2022); ii) the constraints simply require feasibility of the matches.

Furthermore, the solution of the problem in (1) is unique and it translates into a multinomial logit model (McFadden, 1974).[14] In this context, gains from marriage G_{hj} are the utility gains associated to an hj marriage type, with respect to the case in which both spouses remain single, for each hj marriage in each market. Specifically, Choo and Siow (2006b) show that gains satisfy

$$G_{hj} = \log \frac{(\mu_{hj})^2}{\mu_{h0}\cdot\mu_{0j}}. \tag{2}$$

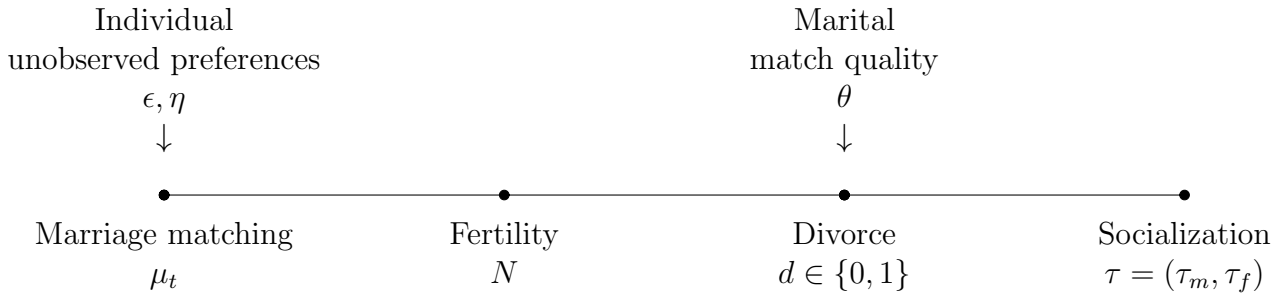
Gains from marriage are therefore exactly identified from the observation of matching patterns μ_t, while individual utilities cannot be without observing transfers across spouses.

3.2 Utility and household choices

In our model, the systematic component of marital utility U_t is endogenous, as we allow for parental investment choices after marriage, inducing a parametrization of individual spouses' preferences which (we show) is over-identified when complementing data on marriage patterns with data on the outcomes of parental investment choices. Marital utility U_{hj} depends on marriage outcomes, determined in turn by future household choices. The timing of the model is illustrated in Figure 2. After households are formed in the marriage market, spouses cooperatively choose fertility, that is, the number of children, N. A match-specific

[14] Galichon and Salanié (2022) and Chiappori et al. (2017), specifically, show that the two-sided matching problem in (1) reduces to a series of one-sided discrete choice problems and characterize the functional form of generalized entropy.

Figure 2: Timing of the Model



quality shock θ is then realized and observed by both spouses. Depending on the realization of this shock, the spouses cooperatively choose $d(\theta)$; that is, they decide whether they remain married ($d = 0$) or they separate ($d = 1$).

Finally, parents invest in the cultural socialization of children. They choose a cultural socialization effort, $\tau(d)$, which increases the probability that children will share the same trait. The choice of $\tau(d)$ is either a cooperative decision by both parents in the household (when $d = 0$) or a non-cooperative decision of the mother in case the household is separated ($d = 1$). 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, in line with the patterns observed in Table C.7. When parents choose socialization, they account for the probability that their children will attain a high or low level of education, recognizing that this probability depends on the education of both parents. To simplify the analysis, we do not explicitly model parents' investment decisions in their children's education; instead, we assume that parents rationally anticipate the probabilities observed in the data.

We now construct the systematic utility of marriage U_t. Recall that the total expected utility of a single male (resp. female) is $U_{h0} + \epsilon_{m0}$ (resp. $U_{0j} + \eta_{0f}$); we assume single individuals have no children. The utility for each married household of type $hj \in C \times C$, has instead an economic and a parental component:

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

The economic component, U_{hj}^{ec}, is the utility deriving from the spouses' economic activities, e.g., from their (present discounted) income. We assume that this component is proportional

to the sum of the utility of an h-type man and a j-type woman would receive as singles:

$$U_{hj}^{ec} = \alpha \left(U_{h0} + U_{0j} \right),$$

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

On the other hand, the parental component U_{hj}^{par} is the indirect utility of the spouses' future choices of fertility, socialization, and divorce. Each child is associated with a cultural-ethnic trait and an education level, a characteristic $k \in C$. The utility component that household hj derives from a child with characteristic k is denoted u_{hj}^{k}. Before the household fertility decision, the characteristic k is a random variable. We now construct its probability distribution to compute the expected utility of a child for a household hj, which we denote u_{hj}. Let P_{hj}^{k} denote the probability that a household $hj \in T$ has a child $k \in C$. We assume the following form of 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 the probability that a household hj has a child of education e. It is exogenous and directly observed in the data. P_{hj}^{c} is instead the probability that an household hj has a child of cultural trait c. It 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 the expected utility u_{hj} as follows:

$$u_{hj} = P_{hj}^{s} \left[\sum_{c \in \{n,i\}} P_{hj}^{c} \left(V_{h}^{cs} + V_{j}^{cs} \right) \right] + P_{hj}^{u} \left[\sum_{c \in \{n,i\}} P_{hj}^{c} \left(V_{h}^{cu} + V_{j}^{cu} \right) \right] \tag{3}$$

Parental preferences over their children's cultural traits and education depend on the parents' culture and education.[15] We restrict the form of such dependence as follows:

$V_{ce}^{c's} = V_{ce}^{c'u} + S_c$, for any $c, c' \in \{n, i\}$; 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, with education e;

$V_{cu}^{c'e} = \gamma_c V_{cs}^{c'e}$, for any $c, c' \in \{n, i\}$; so that γ_c captures the relative strength of the cultural preferences of parents in culture c, with high and low education.

These restrictions imply:

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

and

$$V_{cu}^{c'u} = \gamma_c V_{cs}^{c'u}; \;\; V_{cu}^{c's} = \gamma_c V_{cs}^{c's}.$$

[15] We do not allow, however, individual variation in preferences within cultural and educational group.

Parental utility per child is composed of i) the expected utility deriving from the child's characteristic, u_{hj}; and ii) a direct utility component δ deriving from having a child in the household ($d = 0$) as opposed to outside, i.e., in case spouses choose to separate ($d = 1$). Furthermore, we assume that the marital quality shock θ enters marital utility only if $d = 0$. Finally, we assume that per child utility is proportional to fertility N, and fertility choice entails a cost $\kappa(N)$, increasing and convex in N. We can then formally obtain the parental component of marital utility for household hj as follows:

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

and we proceed to construct it 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 $\tau(d)$ 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:

$$V_{ce}^{ce'} > V_{ce}^{c'e'}, \text{ for all } c \neq c' \in \{n, i\} \text{ and all } e, e' \in \{s, u\}.^{16}$$

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 $c \neq c' \in \{n, i\}$, for given education level e.

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 vertical socialization at the level of the family, we introduce several simplification assumptions. First of all, within a family, all children identify with the same trait. In particular, we abstract from differences in socialization preferences regarding the gender and/or the birth order of children, for recent contributions see Dahl et al. (2022); Felfe et al. (2021). We also abstract from socialization externalities driven by spillover effects across siblings. Secondly, we assume that homogamous native households socialize their children with probability one; 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

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

increasing the probability that the children identify with his trait, c; similarly, the socialization effort of the mother, τ_f, has the objective and the effect of increasing the probability that the children identify with her trait, c'. Crucially, 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. Consequently, the value of the marriage increases in the coordination of investments in children's cultural socialization. Finally, 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:

$$\begin{array}{ll}
P^i_{hj}(\tau,0) = \tau_m + \tau_f + (1-\tau_m-\tau_f)q^i, & P^n_{hj}(\tau,0) = (1-\tau_m-\tau_f)(1-q^i) \text{ if } c = c' = i \\
P^i_{hj}(\tau,0) = \tau_m + (1-\tau_m-\tau_f)q^i, & P^n_{hj}(\tau,0) = \tau_f + (1-\tau_m-\tau_f)(1-q^i) \text{ if } c = i, c' = n \\
P^i_{hj}(\tau,0) = \tau_f + (1-\tau_m-\tau_f)q^i, & P^n_{hj}(\tau,0) = \tau_m + (1-\tau_m-\tau_f)(1-q^i) \text{ if } c = n, c' = i.
\end{array} \tag{5}$$

Assuming that the mother is given custody of children in divorce, socialization probabilities under divorce $P^c_{hj}(\tau, 1)$ are as the ones reported in (5), after imposing $\tau_m = 0$.

The expected utility u_{hj}, as defined by equation (3), depends on the socialization probabilities in (5), 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 entails increasing and convex costs $c(\tau)$. It is then determined as the solution to the following maximization problem:

$$\max_{\tau \geq 0} \; u_{hj}(\tau, d) - c(\tau), \tag{6}$$

Let the solution be denoted $\tau(d)$.[17] 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.

Separation. After observing the realization of the marriage quality shock θ, the spouses optimally choose whether to dissolve the marriage or not, rationally anticipating their total

[17] Whenever possible without confusion, we avoid using the hj subscript in the notation.

utility from the socialization process. Given N, a household of type hj separates, choosing $d(\theta) = 1$, if

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

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

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

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

$$\max_{N} \; N\Big(\pi\left(N\right)u_{hj}(\tau(1),1) + (1-\pi\left(N\right))(\delta + u_{hj}(\tau(0),0))\Big) - \kappa(N). \tag{7}$$

3.3 Discussion

Two of the main modeling choices we have adopted are worth discussing. First, 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 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 their marriage choices. In this case, omitted variable bias is possibly limited once we allow for variation by cultural-ethnic group and education.[18]

[18]Of course the effect of unobservables will be captured in any case if they are idiosyncratic after conditioning on preferences for culture and especially education.

3.4 Results

We describe here informally the most important implications of the model in the previous Section, for a culturally heterogeneous society in which the immigrant group is a minority and the natives the majority.

Socialization. Immigrant parents make costly investments in order to socialize their children, both in culturally homogamous and heterogamous households. Socialization investments in immigrant homogamous households benefit from coordinated incentives. Conversely, socialization investment in heterogamous households depends on cultural intolerance asymmetries. In addition, when married, homogamous households hold a more efficient socialization technology than heterogamous ones. If they separate, the socialization technology is the same, independent of the type of household. As a consequence:

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

For all household types, married or separated, the probability of successful socialization to the trait desired by the parents (or parent) doing the investment is greater than the rate associated with horizontal socialization.[19] We turn now to study comparative statics:

In culturally homogamous immigrant households, whether parents separate or stay married, both parents' socialization efforts are monotonically increasing in cultural intolerance and decreasing in the size of their cultural group. 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 socializes more than the majority, ceteris paribus.[20]

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

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

Separation. Consider a household with positive fertility, $N > 0$. As the utility of marriage derives from socialization, separation leads to a generally less efficient socialization technology:

All types of household stay married if their marriage quality shock is positive; they separate 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 separation:[21]

The separation probability of culturally heterogamous households is higher compared to homogamous immigrant households for the same realization of the stability shock if the mother has higher cultural intolerance. If instead, the father has higher cultural intolerance, the sepaeration probability of culturally heterogamous households is higher compared to homogamous households if and only if the father is an immigrant.

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

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

Fertility. The fertility rates for all types of households are strictly positive, but:

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

Matching. Marital utility exhibits a form of endogenous complementarity in socialization technologies. As a consequence:

The optimal allocation in the marriage market would generate positive assortative matching by cultural-ethnic trait, barring heterogeneity in the idiosyncratic component of preferences and potential market asymmetries in the distribution of cultural-ethnic traits along gender lines in the market.

Education. Highly-educated children increase parental utility in marriage. As a consequence:

[21] Separation choices for heterogamous households 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.

For both culturally homogamous and heterogamous households, independently of parental education, fertility choices are increasing in the parental value of children's education, while cultural socialization effort is unaffected.

The relative strength of the cultural preferences of parents conditional on their education instead affects both socialization and fertility:

Socialization effort and fertility rates of households with (at least one) low-educated spouse are increasing in the relative strength of the cultural preferences conditional on education.

4 Structural estimation: Methodology

We estimate the parameters of the model by observing the marriage patterns, as well as the fertility, separation, and socialization rates. Taking the model in Section 3 to data, we extend it to a larger set of cultural-ethnic groups: Italians, denoted n, and six 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 $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$. The separate marriage markets - one for each of the 20 Italian administrative regions - are denoted $r \in R$. Next, we introduce specific assumptions and functional form parametrizations, describe the parameters of interest, present our estimation procedure, and discuss identification. Table 4 provides a summary of the structural model.

4.1 From the model to the data

Recall that from the model in Section 3, the hj-type marital 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 $\Delta V_{cs}^{c's}$, for all c given education s, and q^c. The economic component of the hj-type marital utility satisfies $U_{hj}^{ec} = \alpha_z \left(U_{h0} + U_{0j}\right)$, where we allow α_z to capture systematic differences between homogamous and heterogamous couples, indexed by $z \in \{het, hom\}$.

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 immigrant 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 zero. Also, we allow socialization and fertility costs to capture systematic differences between homogamous and heterogamous couples, with $z \in \{het, hom\}$, respectively:

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

Table 4: Model: Notation, Variables, and Parameters

Notation	Description
Individual and Household Types	
$c \in \{n, i\}$	Cultural trait: n for native, i for immigrant
$e \in \{s, u\}$	Education level: s skilled (high), u unskilled (low)
$C = \{ns, nu, is, iu\}$	Set of individual characteristics (culture and education)
$h, j \in C$	Male type h and female type j
m_h, f_j	Males m and females f with traits h and j respectively
hj	Household with male h and female j
$h0$, $0j$	Single male h and single female j
T	Set of all household types, including singles
$k \in C$	Child of type k
Choice Variables	
μ_{hj}, μ_{h0}, μ_{0j}	Fraction of marriages hj and singles $h0$, $0j$
N	Fertility: number of children
$\tau = (\tau_m, \tau_f)$	Cultural socialization efforts by parents male (m) and female (f)
$d \in \{0, 1\}$	Divorce decision, where $d = 1$ indicates divorce
Model Parameters	
V_h^k (V_j^k)	Utility of a parent of type h (j) for a child of type k
$\Delta V_{ce}^{c'e} = V_{ce}^{ce} - V_{ce}^{c'e}$	Cultural intolerance of group c with respect to group c', for $c \neq c'$
γ_c	Amplification of cultural intolerance for low-educated parents
S_c	Educational premium for children by culture c
$\sigma_{\tau_z}, \lambda_{\tau_z}$	Socialization cost parameters
$\sigma_{N_z}, \lambda_{N_z}, \xi_z$	Fertility cost parameters
δ	Extra marital value per child (net of socialization)
ρ	Segregation bias in horizontal socialization
ω_{cz}, $\bar{\omega}_{ez}$	Outside options by culture and education
α_z	Economic complementarity in marriage ($z = \text{hom}, \text{het}$)
Utility and Outcomes	
U_{hj}	Total marital utility
U_{hj}^{ec}	Economic utility: $\alpha_z(U_{h0} + U_{0j})$
U_{hj}^{par}	Parental utility, including fertility, divorce, and socialization
G_{hj}	Gains from marriage
π_{hj}	Probability of separation
P_{hj}^k	Probability that household $hj \in T$ has a child $k \in C$
P_{hj}^e	Probability that child has education e (exogenous)
P_{hj}^c	Probability that child has cultural trait c
Idiosyncratic Shocks	
ϵ_{mj}, η_{hf}	Male m and female f idiosyncratic shocks in marriage (also ϵ_{m0} and η_{0f} for singles)
$\theta \sim F(\theta)$	Match-specific marital quality shock

where $\xi_z \geq 1$, captures the dependence of fertility costs on childbearing decisions.[22]

For flexibility in the estimation, we allow the distribution of θ_{hj} to have a mean that depends on the household type hj. 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., $a_{hj} : F(0; a_{hj}, b) = \hat{\pi}_{hj}(0)$.[23] This assumption captures systematic differences in separation rates across household ethnic groups and educational levels without children; that is, independently from children's socialization mechanism. Aggregate evidence is reported in Table C.8.

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 $z \in \{het, hom\}$. This is to indirectly capture differential sorting in both observables and unobservables across cultural-ethnic groups, educational attainment, and between households.

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.[24] A strong network fosters the ability of immigrant 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

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

[23] 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.

[24] 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.6, display the geographical variability in the ethnic groups' distribution across markets.

socialization of the society at large.

4.2 Parameters and moments

The main parameters of interest are the cultural intolerance parameters, $\Delta V_{cs}^{c's}$, for all c, c', given education s.[25] In the estimation, we assume that $V_{cs}^{cs} = V$, for all c; that is, we assume that the value of sharing the same cultural trait as his/her own child is constant for all parents with education s, across all groups c. Furthermore, for identification purposes, we normalize $V = 100$, so that cultural intolerances are measured in units corresponding to percentages of V. With 7 cultural-ethnic groups, we are left with 42 cultural intolerance parameters to estimate. We also estimate education parameters, γ_c and S_c, for all c.[26] 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_z; 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 education 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.[27]

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 (2), where $\hat{\mu}_{hj}$ is obtained from the distribution of marriages over the period 1995-2012, while $\hat{\mu}_{h0}$ and $\hat{\mu}_{0j}$ are taken from the population vectors by ethnic group, education, gender and marital status of individual Census data in 2001 and 2011; see Table C.9 for the distribution of single males and females by culture and education. We compute fertility rates $\hat{N}_{hj}$ as the average number of children in households of type hj, including zeros. 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,

[25] Given our structure of preferences, the intolerance parameters of individuals with high education, $\Delta V_{cs}^{c's}$, are the natural reference. The remaining intolerance parameters are obtained as linear transformations of $\Delta V_{cs}^{c's}$ through γ_c and S_c.

[26] For computational reasons, we restrict γ_n of natives to be equal to the one of Europeans-EU15.

[27] 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 (6) 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.

we construct socialization frequencies, $\hat{P}^k_{hj}$, as the fraction of households of type hj in which children speak a given language k at home.[28] Given normalization restrictions, we end up with a total of 101 parameters to match 7,698 moments.[29]

We estimate model parameters via a method of moments estimator, matching the vector of theoretical moments, $\widetilde{\mathbf{\Pi}}(\boldsymbol{\beta})$, implied by the model for a given vector of parameters $\boldsymbol{\beta}$, to their empirical counterparts, $\widehat{\mathbf{\Pi}}$. Formally, given a weighting matrix $\mathbf{\Omega}$,[30]

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

4.3 Identification

The parameters of our model are well identified. In fact, by incorporating endogenous parental investments into a TU matching framework, we are able to separately identify preferences across spouses. This is a significant advance to the existing marriage matching literature, based on Choo and Siow (2006b). In this literature, in fact, match-specific surplus parameters (i.e., the joint gains from marriage) are exactly identified from equation (2) given the observation of matching patterns; but individual preferences of each spouse cannot be without observing transfers across spouses. However - as we already discussed in Section 3.1 - by embedding parental investments into marriage matching, our model induces a parametrization of individual spouses' preferences which (we show) is over-identified when complementing data on marriage patterns with data on fertility, divorce, and cultural socialization.

Specifically, our estimation procedure fundamentally exploits cross-sectional variation in outcomes across household types by cultural groups and education. To this end, we allow parameters to be cultural-ethnic group specific, but we assume they are constant across households within the cultural-ethnic group and, in particular, constant along gender lines. Furthermore, to be able to also exploit variation in the composition of the population across regions, we assume that all parameters are constant along this dimension and that

[28] 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.

[29] Since q^c is indexed by the region $r \in R$, we obtain a set of moments for each of the 20 regions. 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.

[30] 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 the rest of the 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.

each administrative region corresponds to a separate local marriage market.[31] Finally, our identification strategy hinges on i) random variables θ_{hj} having the same distribution across households hj; on ii) segregation bias, ρ, economic complementarity in marriage, α_z, as well as 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 cost parameters, in particular, implies that any difference in costs across groups and education are bound to be attributed, in our estimates, to cultural intolerances.[32]

Under these assumptions, no restrictions need be imposed on cultural intolerance and education parameters: we separately identify i) the cultural intolerance of parents of culture c with respect to children of culture c', $\Delta V_{cs}^{c's}$, for all cultural groups c, c', and ii) the education preference parameters, γ_c and S_c, for all c. In particular, we can identify the cultural intolerance parameters of minorities with respect to children integrated as natives, ΔV_{is}^{ns}, separately from the cultural intolerances of natives towards all minorities, ΔV_{ns}^{is}, for all immigrant groups i and educational level s. This is, in fact, tantamount to identifying what we have referred to in the Introduction as the demand and supply components of cultural integration as an equilibrium phenomenon. We can also allow the outside options of being single, ω_{c_z} and $\bar{\omega}_{e_z}$, to vary by cultural-ethnic groups, by education, and be specific to each household type.

In Appendix E, we provide a formal proof of identification and Table E.1 summarizes the recursive structure of the identification argument. In the following, we outline the result's logical structure to better construct an intuition for it. Take an immigrant group i. First of all, notice that the socialization moments for homogamous immigrant families where both spouses are educated, P_{isis}^{i}, are independent of education parameters γ_i and S_i and are linear in their cultural intolerance towards the natives, ΔV_{is}^{ns}. The same moments, but for families with at least one uneducated spouse, depend linearly on the product $\gamma_i \Delta V_{is}^{ns}$. Variation across household types by education, therefore, allows us to separately identify γ_i and ΔV_{is}^{ns}.[33] The cultural intolerances ΔV_{is}^{ns} determine the demand for integration on the part of immigrants.

The supply parameters ΔV_{ns}^{is} are instead identified from the socialization moments of heterogamous families. For these households, the equilibrium outcome of parents cooperating

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

[32] Cultural intolerances towards group c might then partly capture the relative cost of learning the language group c speaks as well as the relative value of the language, as captured, e.g., by the size of its global speaking population. This is consistent with our general view that cultural intolerances are related to cultural (and hence linguistic) distances.

[33] More specifically, these moments also depend on the parameters of socialization costs, but these can be identified exploiting divorce and fertility moments; see Appendix E for details.

on the socialization of the children implies that only the parent with higher intolerance socializes, i.e., either τ_m or τ_f is zero. But positive socialization effort adds on pure random socialization, so that, $P^i_{isns} - q^i > 0$ (resp. < 0) would identify the immigrant father (resp. the native mother) as applying positive socialization effort. Furthermore, for these households, socialization effort accounts for the intolerances of both parents, and $P^i_{isns} > q^i$ is linear in $\Delta V^{ns}_{is} - \Delta V^{is}_{ns}$. But since the intolerances on the demand side, ΔV^{ns}_{is}, are identified from socialization in homogamous marriages, the socialization moments of heterogamous households identify the supply side parameters ΔV^{is}_{ns}.[34]

As for the value of children's education, S_c, notice that it affects both separation and fertility choices, and hence, it is identified by divorce and fertility moments. The remaining parameters are identified simply and directly from their corresponding moments: the mean of the marriage quality shock and the value of a child within marriage, a_{hj}, δ, from the divorce moments; the fertility costs from the fertility moments; the outside options to marriage and the relative effects of marriage on spouses' economic opportunities, ω_{c_z}, $\bar{\omega}_{e_z}$ and α_z, from marriage gains, that is, from the average probability of marrying, as in Choo and Siow (2006a). In the proof of identification in Appendix E, we show that these parameters are over-identified just considering the cross-sectional variation across household types by cultural groups and education. Variation across regions simply adds to the over-identification.

5 Structural estimation: Results

This section first describes the model fit, then introduces the parameter estimates, and concludes with validation exercises and several robustness checks.

5.1 Fit of the model

The model fits the data well. Table 5 compares the average observed and predicted socialization and fertility moments - for homogamous and heterogamous households, respectively, and by educational sorting; see also Figure C.7, panel a) and c). The correlation between predicted and observed data is 0.88 for language socialization rates, 0.94 for fertility rates. Moreover, the model is also able to capture i) the general pattern of separation choices across groups, even though separation rates appear to be slightly underestimated; ii) the pattern of marriage gains; see Figure C.7, panel b) and d), and iii) marriage patterns; see Figure C.8.

[34] The caveat in the previous footnote applies to this case as well.

Table 5: Fit of the Model

a) Homogamous households

	Italian Socialization						Fertility					
	High		High-Low		Low		High		High-Low		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
Europe-EU15	0.49	0.47	0.49	0.47	-	-	0.58	1.00	0.72	1.01	0.77	1.03
Other Europe	0.42	0.37	0.44	0.34	0.34	0.33	0.58	0.49	0.64	0.61	0.68	0.71
North Africa-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-Saharan 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 households

	Italian Socialization						Fertility					
	High		High-Low		Low		High		High-Low		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
Europe-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
Other Europe	0.90	0.96	0.96	0.93	0.96	0.90	0.62	0.59	0.54	0.51	0.46	0.49
North Africa-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-Saharan 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, comparing observed (data) and predicted (model) socialization and fertility moments by cultural-ethnic group and education of spouses, for households of highly-educated spouses (high), households of spouses with high and low education (high-low), and households of low educated spouses (low). Estimates are reported for homogamous and heterogamous households in panel a) and b), respectively. Estimates are weighted by the observed number of marriages.

5.2 Parameter estimates

Table 6 presents our parameter estimates.[35] First of all, cultural intolerance parameters are strictly positive in our estimates;[36] that is, parents of each cultural-ethnic group have preferences for socializing children to their own cultural-ethnic group. Secondly, intolerance parameters are highly heterogeneous across groups; that is, some groups are much more resilient in their cultural identity compared to others. Notably, this is the case, e.g., for high-educated immigrants from North Africa-Middle East: A child integrated into the native culture, for a North African parent, is valued 72% less than one socialized to the parent's

[35]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.

[36]Cultural intolerance parameter estimates are statistically significant and standard errors are reported in parentheses in Table 6. Despite their significance, some parameter estimates display larger standard errors due to the unbalanced marriage distribution by cultural and educational attributes of spouses; that is, where natives are a majority and immigrant groups represent small minorities.

Table 6: 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
		(3.003)	(5.371)	(3.161)	(2.997)	(2.994)	(2.816)
$\Delta V^c_{i_E}$, Europe-EU15	16.41		52.60	4.773	88.96	90.81	14.02
	(2.047)		(9.451)	(7.362)	(6.494)	(12.486)	(6.856)
$\Delta V^c_{i_O}$, Other Europe	40.32	23.05		97.14	87.65	89.96	65.01
	(1.272)	(10.416)		(12.453)	(10.803)	(17.314)	(11.564)
$\Delta V^c_{i_M}$, North Africa-Middle East	71.58	7.003	98.21		62.38	60.99	83.59
	(3.816)	(9.487)	(10.341)		(18.509)	(13.134)	(15.007)
$\Delta V^c_{i_A}$, Sub-Saharan Africa	59.25	96.61	83.06	43.08		58.31	54.53
	(2.403)	(9.494)	(9.102)	(8.509)		(10.307)	(17.302)
$\Delta V^c_{i_S}$, East Asia	45.92	46.28	85.83	65.09	34.79		53.81
	(2.254)	(10.132)	(11.589)	(13.894)	(12.956)		(14.626)
$\Delta V^c_{i_L}$, Latin America	31.65	12.18	79.85	91.32	77.51	54.09	
	(2.969)	(6.320)	(10.331)	(14.687)	(23.975)	(13.857)	
			Education Parameters				
Surplus for skilled children, S_c	16.67	11.05	52.95	10.69	1.38	35.52	44.61
	(0.142)	(1.506)	(0.543)	(1.457)	(0.972)	(0.900)	(1.550)
Educational preference of unskilled parents, γ_c	1.05	1.05	1.68	1.12	0.98	1.34	0.96
	(0.001)	(0.001)	(0.024)	(0.025)	(0.0153)	(0.023)	(0.057)
			Cost Function and Extra Parameters				
Socialization Cost Parameters	σ_τ hom	16.985	Fertility Cost Parameters			σ_N hom	65.060
		(2.496)					(1.004)
	λ_τ hom	0.857				λ_N hom	0.117
		(0.030)					(0.032)
	σ_τ het	24.163				ξ hom	1.028
		(9.288)					(0.003)
	λ_τ het	0.451				σ_N het	82.667
		(0.380)					(0.801)
Extra Marital Gain per Child	δ	0.848				λ_N het	0.195
		(0.142)					(0.028)
Segregation Parameter	ρ	1.474				ξ het	1.017
		(1.161)					(0.012)
Economic complementarity in marriage	α hom	0.425					
		(0.017)					
	α het	0.427					
		(0.039)					

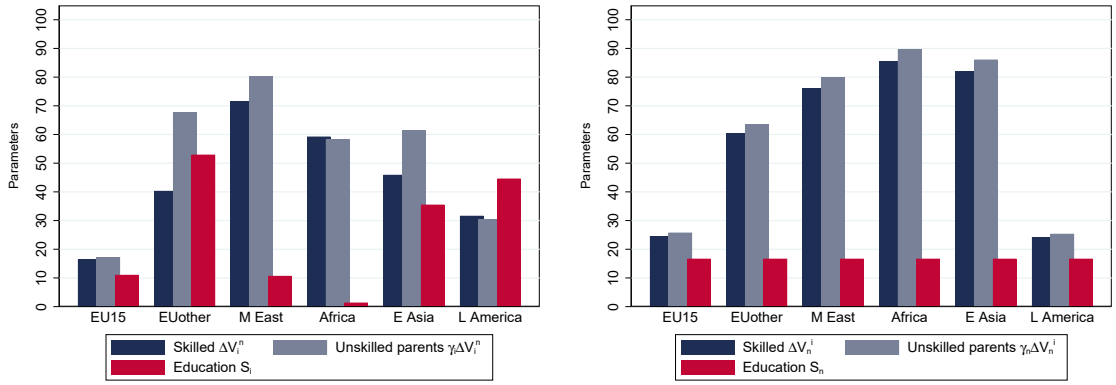
Notes: This Table shows the estimates of the structural parameters of the model. Standard errors are reported in parentheses.

culture. For a highly educated parent of an EU15 country, this loss is only about 16%. The cultural intolerances of Italian natives are also positive and heterogeneous towards different minorities. To an Italian parent, a child socialized to the Sub-Saharan African cultural trait implies a 85% loss, a much larger loss than if the child were socialized to the Latin American trait, about 24%. Thirdly, the γ_c parameters are estimated close to one or even larger than one; that is, cultural intolerances tend to be equal or higher for immigrant parents with low rather than high education. This is the case, especially for Other European and East Asian parents; see Figure 3 (grey bars). Finally, the matrix of intolerance parameters is largely asymmetric; that is, the intolerance of group c versus group c' is often not quantitatively close to the intolerance of group c' versus group c; see Figure 3 (blue bars). Notably, e.g., natives appear particularly accepting of Latin American immigrants, as we already noted;

Figure 3: Cultural Intolerance and Education 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} and natives versus immigrants ΔV_{ns}^{is}, in panel a) and b), for natives n and all cultural-ethnic minorities i, separately for parents with high (blue) and low (grey) education, as well as education parameters S_i and S_n (red).

while Latin Americans are less so towards natives.

The differences in cultural intolerance preferences across groups also translate into significant differences in socialization efforts. For instance, the direct socialization effort of culturally homogamous households, τ, calculated in the extreme case in which the family belongs to a full minority i ($q^i = 0$), ranges between 0.52 (Europe-EU15) and 0.67 (North Africa-Middle East); see Figure C.9.[37]

Concerning our estimates of preferences for children's education, two interesting results are worth highlighting; see Figure 3 (red bars). First, as for cultural intolerance, the value of children's education, S_c, is heterogeneous across cultural groups. It is higher, for instance, for Other European, Latin American, and East Asian minorities, in order. Second, the value of children's education is generally dominated by cultural preferences but with substantial heterogeneity: the ratio between S_c and the cultural intolerance ΔV_{is}^{cs} ranges between about 67 percent for Europeans-EU15 and only 2 percent for Sub-Saharan Africans. These differences imply a large variation across cultural-ethnic groups in their willingness to trade off cultural distance for education in children.

Our estimates show that socialization costs across households are significantly different. The cost σ_τ of heterogamous families is about twice as high as the one of homogamous families.[38] This difference in socialization costs reinforces the gap in socialization investments

[37] Recall that socialization efforts are independent of the additive value of education S_c.

[38] Costs functions are assumed to be independent of the cultural-ethnic group of spouses. The estimates of λ_τ, associated with the degree of convexity of costs, are comparable across family types.

between family types. We also estimate a positive difference in fertility cost σ_N, one-third greater for heterogamous than for homogamous families.Overall, our estimates imply that fertility investments are much more costly whenever spouses belong to different cultural groups. We also 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 50 percent greater than the contribution implied by their actual representation in the population under random matching. This bias parameter indirectly measures 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 households (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, estimates are reported in Table C.10.

5.3 Validation and robustness

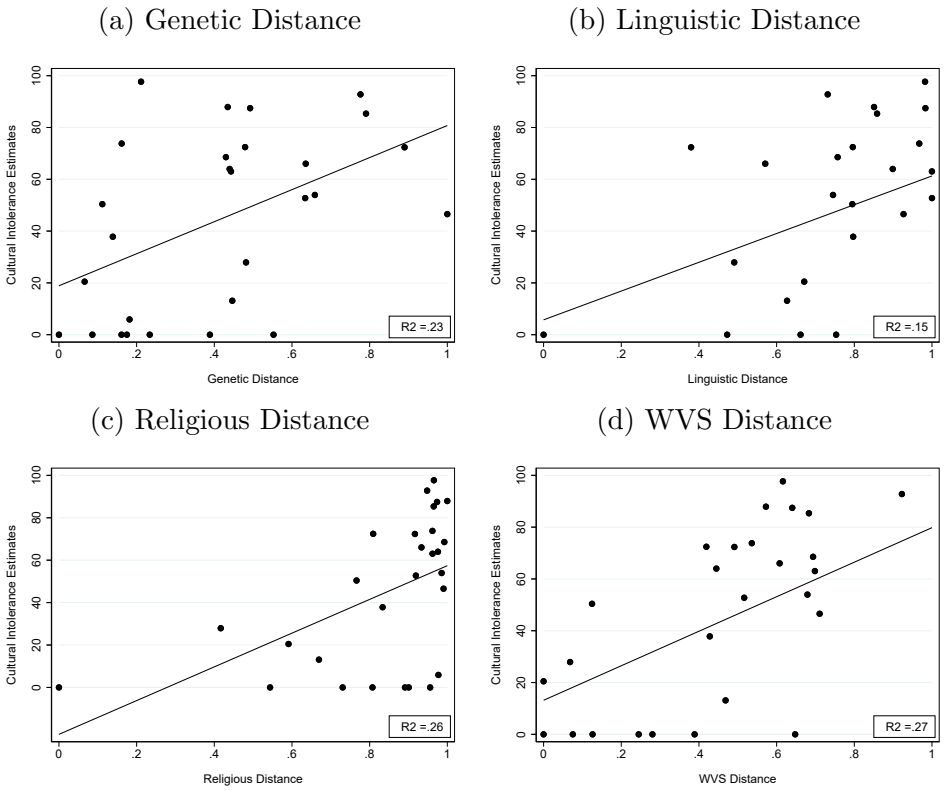
We present, in turn, several different exercises to validate our estimates. We first externally validate our cultural intolerance estimates across cultural-ethnic groups by comparing them to common cultural distance measures (Spolaore and Wacziarg, 2009, 2016), as shown in Figure 4. These measures explain 15 to 30 percent of the variation in cultural intolerance (40 to 70 if we weight estimates by the number of marriages per match). This is remarkable as cultural distance measures are, by construction, symmetric, while our estimates impose no symmetry restrictions. Furthermore, we validate our estimates on out-of-sample data; specifically: i) marriages of newly formed couples (2013-2019) - see Figure C.10; ii) socialization rates of divorced couples - see Figure C.11; iii) pattern of marriages at the provincial level, a higher geographical disaggregation level than our main exercise - see Figure C.12.[39]

Regarding robustness, we present our main exercise here, while additional robustness checks are reported in Appendix D. Our main exercise refers to our proxy for cultural-ethnic socialization, *language spoken at home*. Ethnic identity and spoken language are relevant culturally related specific attributes, and both allow the direct transmission of cultural characteristics across generations.[40] While, reasonably, all children living in Italy learn Italian at school, reporting to speak Italian at home - when at least one spouse is an immigrant - in our interpretation, reveals that children are well-integrated into the native Italian cultural-ethnic identity (Bazzi et al., 2019; Salari, 2020). To corroborate this interpretation, we show that

[39] Due to data limitations, these exercises focus on marriage distribution by cultural-ethnic lines, excluding educational variability; see Table C.11.

[40] 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.

Figure 4: Cultural Intolerance Estimates and Cultural Distance Measures



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

our measure of Italian linguistic socialization influences both the achievement and educational choices of immigrant students. Using student-level data on standardized test scores in reading and math administered by INVALSI,[41] we find that immigrant students who speak Italian at home perform significantly better by a 0.20 (0.11) standard deviations in reading (math), see Table 7. Column (3) of the same table highlights the language's long-term impact on educational careers: speaking Italian at home drives students into high-demanding schools - academic or technical high schools (as opposed to vocational ones). Further supporting evidence is presented in Table C.12, where we document that children who speak Italian at home exhibit larger social integration networks, higher educational achievement, and greater proficiency in the Italian language. Results are based on the *Condition and Social Integration of Foreign Nationals Survey, 2011-2012* survey.

Appendix D presents three additional robustness checks addressing the possibility that:

[41]The test is administered to all students in Italy at the end of grade 5, it is identical for all students, and it is blindly scored; hence results are fully comparable across schools. We describe in detail the Italian educational setting and INVALSI data in Appendix B.

Table 7: 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 and math standardized test scores 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 (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), household 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 the school level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1. Results are robust to various checks; see Table C.13.

(i) fertility differences reflect systematic variation in spouses' observables; and (ii) population distribution across regions may be endogenous; iii) educational attainment of children depends on the marriage status of their parents; that is, whether they are divorced.

6 Integration patterns

In this section, we simulate the dynamics of the distribution of cultural-ethnic traits in the population induced by our structural model.[42] We first present the inter-generational dynamics of cultural traits and education in the population. We then introduce a counterfactual exercise, studying the effects of changes in cultural intolerance parameters on the dynamics.

[42]This exercise rests on the assumption that parameters are invariant over time. Preference parameters - like cultural intolerances - may instead change over time, even relatively rapidly in some instances e.g., if stereotypes or norms of behavior determine them. Consequently, rather than as predictions, the simulations in this section should then be interpreted to highlight the implications of our estimated model. Indeed, we would need a repeated inter-generational cross-section of marriage data to identify the dynamics of cultural intolerances. Also, the notion of cultural integration we adopt continues to refer to 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. Simulations about the dynamics of intermarriage over successive generations, which can be interpreted as a different form of integration, are reported in Figure C.13.

6.1 Inter-generational 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 population distribution by region and ethnic group in our data.[43] 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 of the population of the children of the population at time t, p_{t+1}.[44] The same procedure recursively induces $p_{t+2}, p_{t+3}, ...$[45]

The simulated inter-generational 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 5, normalized so that $q_t^i = 1$ in $t = 0$ for comparability.[46] Although our 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. More precisely, let the *cultural integration rate* of a specific cultural-ethnic group be defined as the reduction of this group as a fraction of the total population (immigrants and natives).[47] According to this definition, we simulate a substantial 75% integration rate overall, in a single generation. By the third generation, 97% of immigrants are simulated to have integrated into the native Italian culture. The rate of integration is however heterogeneous across cultural-ethnic groups. In particular, the North African-Middle Eastern minority integrates almost completely, 92%, in a single generation. A similarly fast pattern of convergence is also displayed by the European-EU15, East Asian, and Other European minorities. On the other hand, a significantly slower integration rate characterizes the Latin American minority, with an integration rate of 54%, which is simulated to reach full convergence only by the fourth

[43] 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.

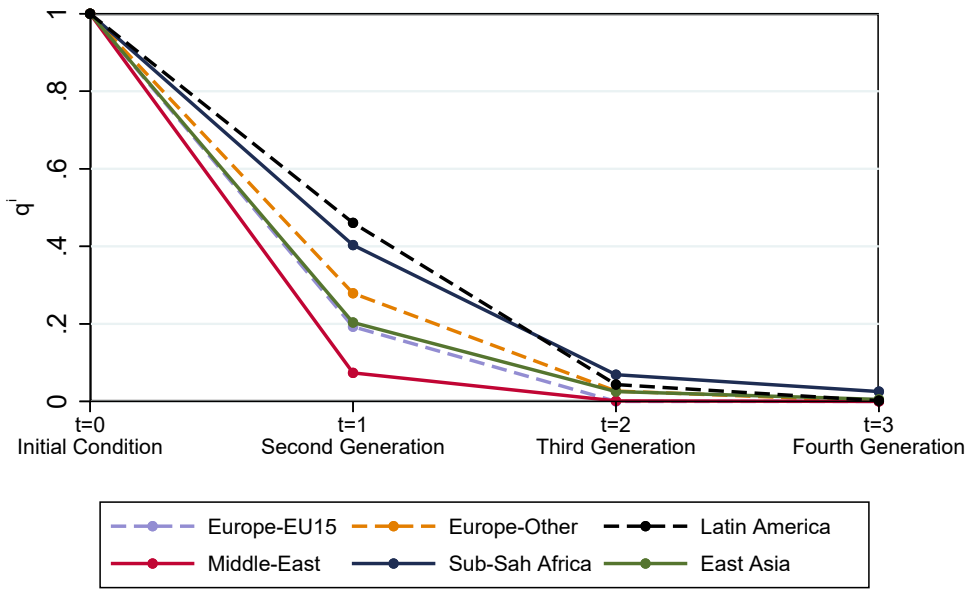
[44] Reproduction is asexual in the model, hence we consider future generations populated by men and women of equal proportion. Note also that the individuals in the population composing the distribution p_t are distributed across the age dimension. We disregard this in the estimates, and hence also in the simulation.

[45] At each step, we compute the equilibrium in the marriage market subject to feasibility constraints, represented by equation (1). To handle the large dimensionality of the problem, we take advantage of the iterative projection fitting procedure (IPFP); see Galichon and Salanié (2022); Galichon (2018).

[46] 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. See Figure C.14 for the inter-generational dynamics of cultural traits without normalization.

[47] 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.

Figure 5: Inter-generational Dynamics of Cultural Traits

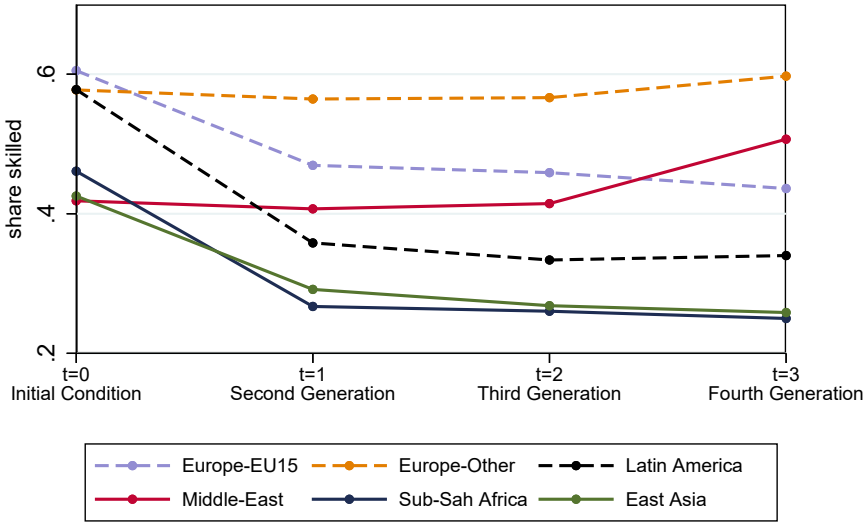


Notes: This Figure shows the inter-generational dynamics of the distribution of cultural traits in the population for minority groups (each with population normalized to 1 in $t = 0$), over successive generations.

generation. A slower integration rate also characterizes the Sub-Saharan African minorities, 60% in one generation.

These heterogeneous integration patterns document various interesting different mechanisms operating in the model. The patterns of cultural integration of European-EU15 and Other European minorities are the direct result of their relatively low cultural intolerance. On the other hand, the East Asian and Sub-Saharan African minorities' slower integration is due in part to their higher intolerance. But cultural intolerance is not the only determinant of the dynamics of integration of different cultural-ethnic groups. Homogamous marriage rates, fertility rates, and other demographic characteristics turn out to have sizable independent effects on cultural integration in the simulations. This is clearly illustrated by the fact that, while North Africa-Middle East, Sub-Saharan Africa, and East Asia show relatively comparable cultural intolerance, they display significant differences in the dynamics of integration. Indeed, a strong estimated selection into homogamous marriages of immigrants 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, since natives display low cultural intolerance towards them, they turn out to be uniquely able to socialize children also

Figure 6: Inter-generational Dynamics of Education: Share of Highly-Educated Individuals



Notes: This Figure shows the inter-generational dynamics of the distribution of highly-educated individuals by cultural-ethnic group, i.e., conditional on having at least one parent belonging to that cultural-ethnic group.

in heterogamous marriages with natives.

The simulated pattern of cultural integration depends on the educational standing of the households: on average we simulate a slower pattern of integration for immigrant individuals with low education; see Figure C.15 for the inter-generational dynamics of the fraction of the population by cultural-ethnic trait and by education. This gap in convergence is in part explained by the sizable differences in selection into homogamous marriages by education, due in turn to the relatively higher cultural intolerances of immigrants with low education.

The inter-generational dynamics of immigrants' education levels can be considered an indication of the economic integration of immigrant groups. Figure 6 reports the evolution of the share of individuals with high education in the population, by cultural-ethnic group of their generation $t = 0$ ancestor. These dynamics, of course, depend on the probability of children becoming highly-educated for each household type, but they are also affected by integration rates, as natives have highly-educated children with relatively higher probability. The initial drop in education levels, for all groups, is just a consequence of data misalignment.[48] Figure 6 therefore essentially displays constant immigrants' education levels over successive generations. The only significant exception is represented by North African and

[48]Because of data limitations, the distribution of the population by education of generation $t = 0$ is constructed classifying as highly-educated all individuals with a high school degree, while for second-generations onwards, the same classification is based on high-track high school degrees, as opposed to vocational schools degree; see Appendix B for details.

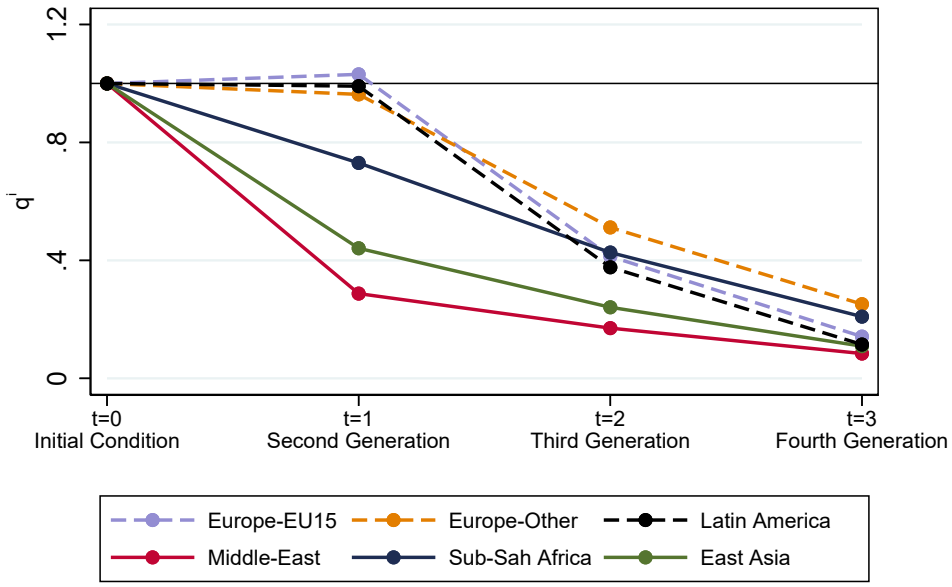
Middle-Eastern minorities, which slowly increase their educational attainment over time, in part as a consequence of a relatively high cultural integration rate.

6.2 Counterfactual cultural intolerance parameters

In this section, 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 $\Delta V_{cs}^{c's}$, for c, c'.

Consider 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, the counterfactual environment in which natives are fully tolerant towards the immigrants' cultural diversity. In this counterfactual simulation, the tolerance of natives increases the values of intermarriages for both natives themselves and immigrants, who now can socialize their children to their own cultural identity even when married to a native. The simulation shows a large increase in intermarriages with natives coupled with a substantial increase in the fertility of these marriages and a decrease in socialization to Italian. All these effects induce a reduction in immigrants' integration compared to the baseline. The dynamics of integration are still heterogeneous across groups, however. 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; see Figure 7 and Figures C.16-C.17

Figure 7: Inter-generational Dynamics of Cultural Traits with Italians Fully Tolerant



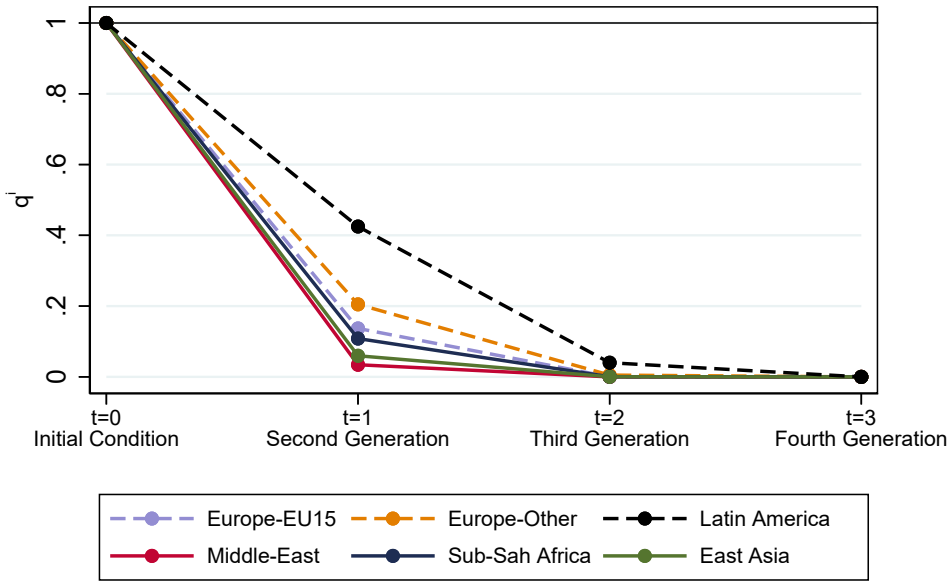
Notes: This Figure shows the inter-generational dynamics of the distribution of cultural traits in the population for minorities (with population normalized to 1 in $t = 0$), over successive generations assuming the case of complete tolerance of the Italian majority towards minorities, $\Delta V_{ns}^{is} = 0$.

for the mechanisms driving the dynamics of integration.[49]

Fouka (2020) empirically studies the effects of a shock of opposite sign, an increase in V_{ns}^{is} reflecting a lower supply of acceptance from natives, driven by restrictions on horizontal socialization. Specifically, she examines the rise in anti-immigrant sentiment in the U.S. after WWI, which led to language bans in elementary schools. She finds higher homogamy rates and strengthened vertical socialization (measured by the prevalence of children's foreign names, a proxy closely related to our measure of speaking the language of origin at home). This is consistent with the implications of our model. Interestingly (and counterintuitively), however, when in our simulations we account for all behavioral effects through the whole marriage market, fertility, and divorce, this kind of shock induces higher rather than lower integration rates in equilibrium.

Consider instead the case of a an increase in the cultural intolerance parameters of immigrants ΔV_{is}^{ns}, interpreted e.g., as a reduction in the job opportunities available for the immigrants.[50] When we simulate the population dynamics in this case, we obtain the somewhat

Figure 8: Inter-generational Dynamics of Cultural Traits net of Economic Incentives



Notes: This Figure shows the inter-generational dynamics of the distribution of cultural traits in the population for minorities (with population normalized to 1 in $t = 0$), given a proportional increase in ΔV_{is}^{ns} for all minorities by 20% of ΔV_{ns}^{is}.

[49]In the opposite case, in which Italians are fully intolerant towards all minorities, the dynamics of integration follow the results in the baseline, with an integration rate of about 80% overall over a single generation; see Figure C.18.

[50]Since job opportunities are generally in control of the native population, we postulate that such an increase in ΔV_{is}^{ns} would be, for any immigrant group i, proportional to the natives' cultural intolerance for that group. In particular, in this exercise, we increase ΔV_{is}^{ns} by 20% of ΔV_{ns}^{is}.

counter-intuitive result that the integration of minorities towards Italian culture accelerates by about 10 percentage points compared to the baseline, where immigrants have higher economic incentives to integrate; see Figure 8. In fact, with lower economic incentives for integration, immigrant parents care more about socializing their children, and hence the value of both homogamous and heterogamous marriage significantly reduces for immigrants (while the value of homogamous marriages of natives remains unchanged) because the stronger attachment to their identities makes marriage a riskier and costlier investment. This leads to a general equilibrium effect compressing the marriage market for immigrants and reducing their fertility. 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 - because of intense parental socialization - but the fraction of the total population (immigrants and natives) integrating into the native culture, on the contrary, is higher. These results are consistent with Adda et al. (2025), who study how immigrants' marriage and separation decisions respond to a change in the economic value of marriage. Following the 2004 and 2007 EU enlargements, which granted legal status and work rights to Eastern European migrants, intermarriage rates declined, and separation rates among these intermarriages rose. In our model, this reflects a reduced economic incentive to integrate (a larger ΔV_{is}^{ns}), which translates into the effects observed in Adda et al. (2025).

7 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 immigrant integration over time respond to variations both in the demand of immigrants for the preservation of their cultural identity as well as in the 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. For instance, we show that a reduction in the economic incentives for cultural integration might accelerate the immigrants' integration rate. These conclusions might represent a starting point for a debate about immigration policies, with far-reaching implications for societies becoming more ethnically heterogeneous.

References

Abdelgadir, A. and V. Fouka (2020). Political Secularism and Muslim Integration in the West: Assessing the Effects of the French Headscarf Ban. *American Political Science Review 114*(3), 707–723.

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 (2025). There's More to Marriage Than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations. *Journal of Political Economy 133*(4), 1276–1333.

Ahn, S. Y. (2024). Matching Across Markets: Theory and Evidence on Cross-Border Marriage. *Journal of Labor Economics, Forthcoming*.

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), 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.

Caiumi, A. and G. Peri (2024). Immigration's effect on us wages and employment redux. Technical report, National Bureau of Economic Research.

Campo, F., S. Giunti, M. Mendola, and G. Tura (2023). Political Backlash to Refugee Settlement: Cultural and Economic Drivers. Technical report, IZA Discussion Papers.

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 (2022). Goals and Gaps: Educational Careers of Immigrant Children. *Econometrica 90*(1), 1–29.

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., D. L. Nguyen, and B. Salanié (2019). Matching with Random Components: Simulations. *Working Paper, Columbia University*.

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 (2006a). Estimating a Marriage Matching Model with Spillover Effects. *Demography 43*(3), 463–490.

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

Ciscato, E. and S. Weber (2020). The Role of Evolving Marital Preferences in Growing Income Inequality. *Journal of Population Economics 33*(1), 307–347.

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.

Dahl, G. B., C. Felfe, P. Frijters, and H. Rainer (2022). Caught between Cultures: Unintended Consequences of Improving Opportunity for Immigrant Girls. *The Review of Economic Studies 89*(5), 2491–2528.

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. (1997). The Effects of Education, Parental Background and Ethnic Concentration on Language. *Quarterly Review of Economics and Finance 37*, 245–262.

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.

Felfe, C., M. G. Kocher, H. Rainer, J. Saurer, and T. Siedler (2021). More opportunity, more cooperation? The behavioral effects of birthright citizenship on immigrant youth. *Journal of Public Economics 200*, 104448.

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 Papers and Proceedings 107*(5), 251–55.

Galichon, A. and B. Salanié (2022). Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *The Review of Economic Studies 89*(5), 2600–2629.

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.

Gualdani, C. and S. Sinha (2023). Partial Identification in Matching Models for the Marriage Market. *Journal of Political Economy 131*(5), 1109–1171.

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 Elémentaires 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.

Shapley, L. S. and M. Shubik (1971). The Assignment Game I: The Core. *International Journal of Game Theory 1*(1), 111–130.

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.

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.

Appendices For Online Publication

A Data and Sample Construction

We obtained restricted access to administrative Italian data at the individual level from ISTAT through its Laboratory for Elementary Data Analysis (ADELE).[51] 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 the 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. 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 Figure C.3 for the bivariate marriage distribution. Crucially, these data account only for legal marriages and only when celebrated in Italy.[52]

[51] 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.

[52] 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 relevant for our study since they are typically less stable and hence less

Fertility. Fertility rates come from municipality birth registries, which contain the universe of individual birth records of residents in the municipality of enrollment, 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. Unfortunately, these data do not provide information on the educational level of spouses.

Fertility from Census 2001 and 2011. We recover fertility by cultural-ethnic group and, importantly, spouses' education from the decennial Census of 2001 and 2011, which contain the universe of the population resident in Italy in 2001 and 2011. Individual Census data provide information about the family and the family identifier, the municipality of residence, and the main socio-demographic characteristics of individuals including gender, age, marital status, and education. For immigrant individuals, Census data provide additional information about the nationality, the country of origin, and the year of arrival in Italy. We identify marriages as heterosexual love relations between spouses who are living together and who are legally married, by exploiting information on the family id, marital status, and type of the relationship. Coherently with our main dataset, we focus on marriages formed after 1995 for Census 2001. Census 2011, instead, does not report the year of marriage. We thus select marriages of *natives* based on spouses' years of birth, i.e., we select native men born from 1955 to 1984 and native women born from 1959 to 1988. In addition, we select marriages of *immigrants* with children born after the year of arrival in Italy and we select marriages of immigrants not arriving in the same year and from the same country, as likely marriages formed outside the Italian marriage market. Even applying these selections, in this sample, fertility rates are mechanically higher compared to our main analysis. Thus, from these data, we exploit the variation in fertility by spouses' education and we calculate the fertility rate (number of children including zeros) by cultural-ethnic group and education of spouses. We then match these fertility rates with the average fertility rates derived from municipality birth registries to recover variation in fertility choices by spouses' education.

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.[53] We

motivated by fertility and children socialization; see (Lundberg and Pollak, 1993; Chiappori et al., 2017). Also, marriages celebrated outside of the Italian marriage market are arguably less motivated by cultural integration issues.

[53]For our investigation period, registries of civil court chancelleries constitute the unique source for separations data, while starting from December 2014 (in the application of Law n. 162/2014) consensual separation

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 behaviors. The data allow us to analyze various aspects of the marital dissolution phenomenon. We investigate, in particular, the custody assignment of children.[54]

Single Individuals. We derive the population vectors by ethnic group, education, gender, and marital status from individual Italian Census data of 2001 and 2011. We select adult individuals (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 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. The distribution of single male and female adult individuals is reported in Table C.9.

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 households. The survey targeted foreign residents in Italy and it was conducted at the household level to provide socio-demographic information about all household members,

proceedings can be submitted to the civic registrar. This rules out potential sample selection concerns.

[54] 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 but small differences in custody assignment conditional on mother and father migrant status. We abstract from incorporating them in the model for the sake of simplicity.

for a total sample of 25,356 respondents.[55] 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.[56] 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.[57]

We exclude from our analysis, single respondents and households without children. For our analysis, we consider children and young adults 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 households, 86.7% of these households 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. 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 household information across different sources. We matched marriage, birth, and separation records on the exact date of marriage and spouses' 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 99.5% of marriages, and we discard the remaining fraction. Such low percentages suggest that marriages can be uniquely

[55] 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.

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

[57] 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?*

identified through the set of time-invariant characteristics listed above. Unfortunately, during the revision process, ISTAT changed data availability rules for researchers and we are no longer able to match marriage and fertility data. We then matched fertility moments by cultural-ethnic group and education of spouses, and region as described above. The final sample of 4.2 million marriages corresponds to 94% of the universe of marriages celebrated in Italy during the time interval 1995-2012. In the final dataset, the fertility rate corresponds to 69.6% 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 (2), 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, μ_{h0}, μ_{0j}. Individuals are heterogeneous both in their cultural-ethnic traits and their education level.

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

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

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

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, $d = 0$ as follow (for all regions r):

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

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.

B Education in Italy and INVALSI data

From the ISTAT data on marriages, the education levels of spouses can only be constructed by classifying as highly-educated individuals with a high school degree. For second-generations onwards, however, we have obtained administrative student-level data which allows us to construct a finer classification. These data provide information 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).[58] 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 into 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. (2022), 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., 2022).

B.2 INVALSI tests and questionnaire

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

[58] 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/.

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-ended questions, where the exact structure varies by grade.[59]

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 information from a student's questionnaire, including a specific question on the main language spoken at home by students, similar to our main analysis. The questionnaire is administered only in grade 5 (the last year of elementary school) and in grade 10 (the second year of high school).

B.3 Sample selection

In our robustness analysis in Section 5.3, 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 career 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 low-educated

We exploit INVALSI data to recover the probability that children become low-educated, i.e., enroll into *low-track* vocational schools compared to high-track schools, conditional on parents' cultural-ethnic traits and educational background, by region. We identify native and immigrant parents thanks to the country of origin, and for immigrant parents, we distinguish three cultural groups: European-EU15, Other European countries, and all other immigrant minorities are grouped together. The first two cultural groups exactly coincide with our cultural-group classification. The last group, instead, aggregates immigrants coming from very different origins and potentially different educational backgrounds. Thus, we recover

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

variation in the probability that children become low-educated across immigrant minorities by exploiting the ISTAT data from the *Condition and Social Integration of Foreign Nationals* Survey, described above. Table B.1 reports summary statistics of the average fraction of children with low education by ethnic group and education of spouses. Descriptively, estimates suggest that children of homogamous marriages of immigrants are less likely to enroll into high-track schools, especially so for children originating from Sub-Saharan Africa, followed by North Africa-Middle East and East Asia. Moreover, as expected, students with low-educated parents are less likely to enroll into high-track schools (columns 5 and 6), compared to students with highly-educated parents (columns 1 and 2).

As a validation exercise, in Table B.2, we focus on Other European immigrants, the cultural group that coincides in the two datasets, INVALSI and ISTAT survey. We show that the average fraction of Other-European children with low education - by education of spouses - is similar when computed from INVALSI data as in panel a), or from ISTAT survey data as in panel b).

Table B.1: Fraction of "Low educated" Children by Ethnic Group and Education of Spouses

	(1)	(2)	(3)	(4)	(5)	(6)
	High-high		High and Low		Low-low	
	Homog	Heterog	Homog	Heterog	Homog	Heterog
Italian	0.29	0.39	0.50	0.56	0.65	0.66
Europe-EU15	0.55	0.33	0.64	0.52	0.71	0.64
Other Europe	0.56	0.34	0.66	0.53	0.76	0.63
North Africa-Middle East	**0.77**	0.33	**0.82**	0.52	**0.86**	0.63
Sub-Saharan Africa	**0.91**	0.58	**0.81**	0.68	**0.83**	0.76
East Asia	**0.77**	0.59	**0.85**	0.67	**0.83**	0.76
Latin America	**0.72**	0.58	**0.83**	0.68	**0.88**	0.77

Notes: This Table reports the average fraction of children with low education by ethnic group of spouses. The outcome variable is an indicator of whether the child enrolls in vocational schools compared to academic and technical schools. Estimates are reported separately for homogamous (odd columns) and heterogamous (even columns) households by spouses' education, distinguishing households with both highly-educated spouses (columns 1 and 2), households with one highly-educated spouse (columns 3 and 4), and households of both spouses with low education (columns 5 and 6). Sources: ISTAT survey data (estimates in bold); INVALSI data (other estimates for homogamous families).

Table B.2: Validation: Fraction of "Low educated" Children for Other European

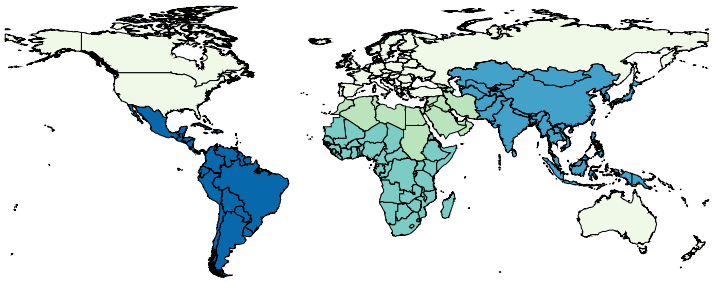
	(1) Mean	(2) High Educated	(3) High-Low Educated	(4) Low Educated
		Panel a) INVALSI		
Prob of low education	0.690	0.593	0.684	0.778
	[0.648,0.731]	[0.543,0.643]	[0.650,0.718]	[0.740,0.817]
		Panel b) ISTAT Survey 2011		
Prob of low education	0.752	0.592	0.812	0.831
	[0.542,0.962]	[–]	[0.788,0.835]	[–]

Notes: This Table reports the average fraction of children with low education for culturally homogamous Other European households, computed on INVALSI data (panel a) and ISTAT survey data (panel b). The outcome variable is an indicator of whether the child enrolls in vocational schools compared to academic and technical schools. Estimates are reported separately for households by spouses' education, distinguishing households with both highly-educated spouses (column 2), households with one highly-educated spouse (column 3), and households of both spouses with low education (column 4).

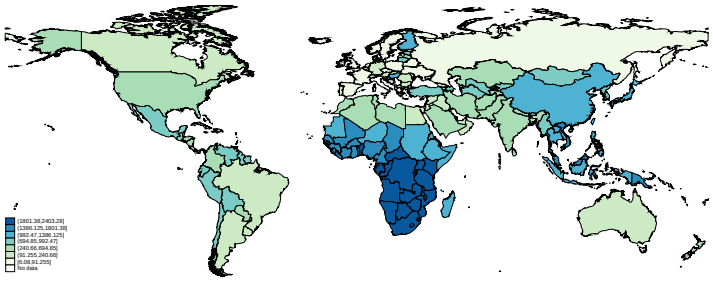
C Additional Figures and Tables

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

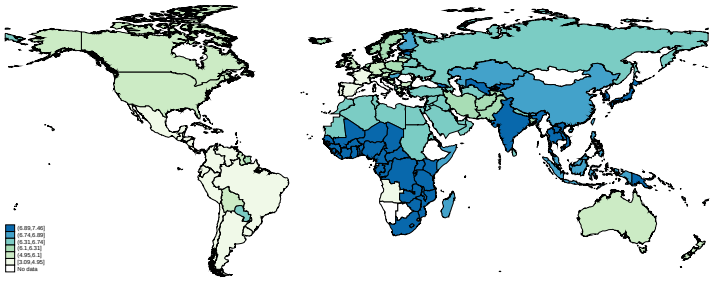
(a) Our Cultural-Ethnic Group Classification



(b) Genetic Distance to Italy



(c) Linguistic Distance to Italy



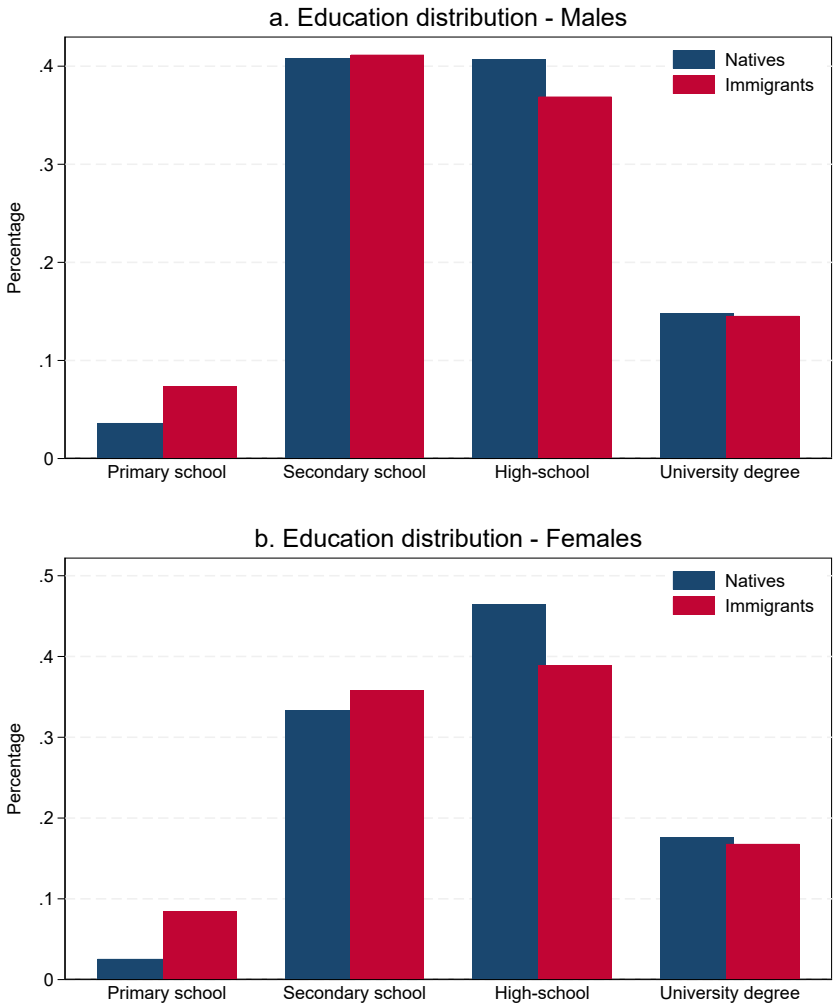
Notes: This Figure shows our classification of countries in cultural-ethnic groups (panel a) and plots the cultural distance of each country towards Italy as proxied by genetic (panel b) and ethnolinguistic distance (panel c). Data for genetic and ethnolinguistic distance are available thanks to Spolaore 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

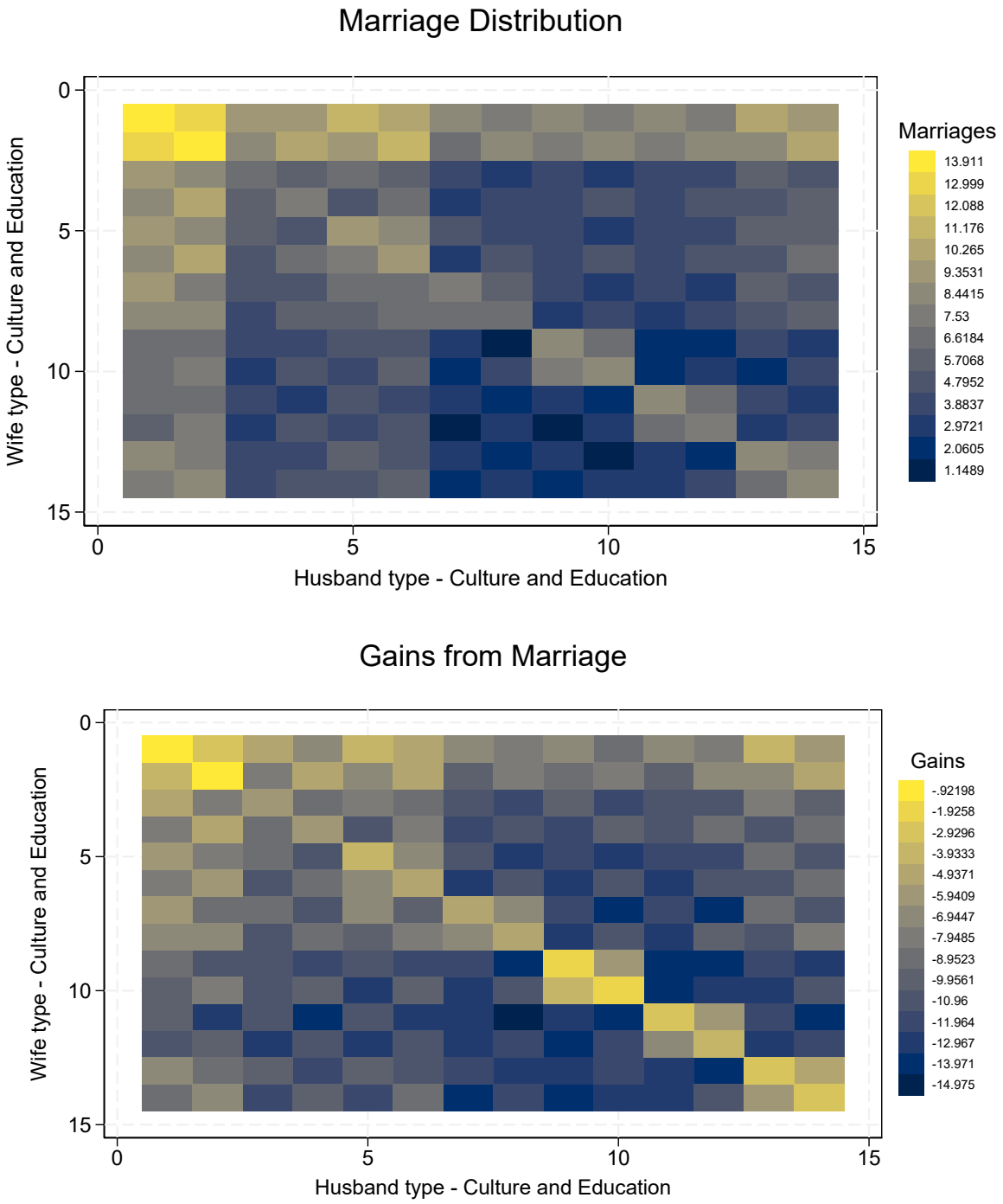
Notes: This Table reports our classification of foreign countries by cultural-ethnic group. We note that the Other European group includes Eastern European countries which became EU members after the enlargements in 2004 and 2007. These enlargements altered the incentives of some immigrants to marry natives, hence the composition of intermarriages in the data is an average over the period; see Adda et al. (2025).

Figure C.2: Education Distribution for Native and Immigrants



Notes: This Figure shows the distribution of educational attainment (in percentages) for natives and immigrants, separately for males and females in panels a and b, respectively. Educational attainment is divided into four categories: (i) primary school degree, (ii) secondary school degree, (iii) high school degree, and (iv) university degree, which includes bachelor's, master's, and PhD degrees. The figure shows that education levels are concentrated around intermediate values (lower and upper secondary), while the tails –especially the lower end (primary school only)– have limited support.

Figure C.3: Marriage and Gains from Marriage Distribution (Heatmaps)



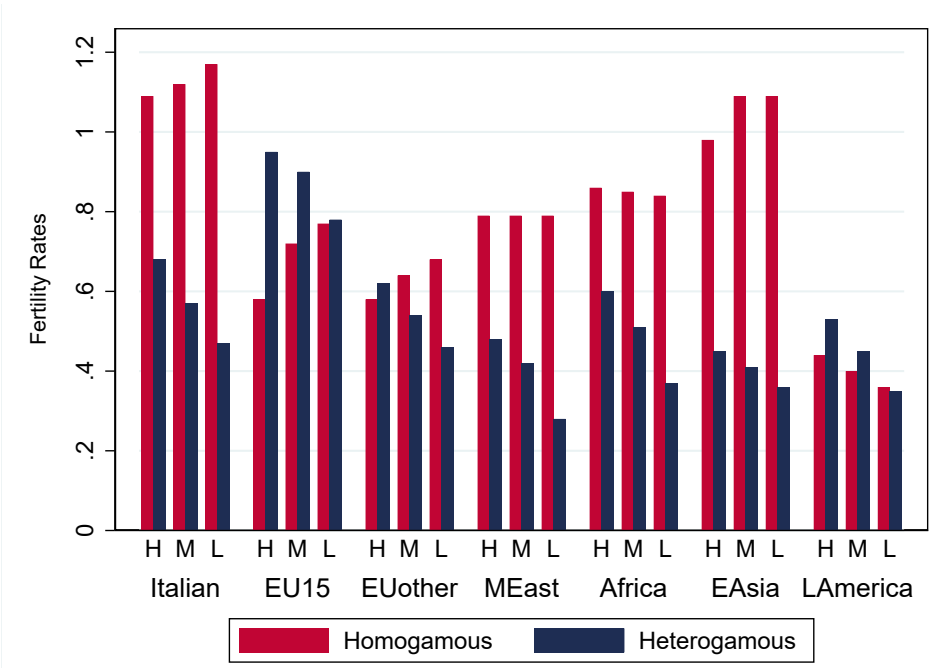
Notes: This Figure reports the heatmap of the bivariate marriage distribution (in log) and the distribution of gains from marriage by cultural-ethnic group and education of spouses. We consider seven cultural-ethnic groups: Italian, European-EU15, Other European, North African and Middle-Eastern, Sub-Saharan African, East Asian, and Latin American; and we distinguish two levels of education: highly-educated individuals and low educated individuals.

Table C.2: 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
R-squared	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 (2), 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) fixed effects. Robust standard errors clustered at the regional level in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Figure C.4: Fertility Rates by Ethnic Group and Education of Spouses



Notes: This Figure shows average fertility rates by ethnic group and education of spouses (H: both highly-educated spouses, M: high and low-educated spouses, L: both spouses with low education). Estimates are reported separately for homogamous (red) and heterogamous (blue) households.

Table C.3: Fertility Rate by Culture and Education

	(1)	(2)	(3)	(4)
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		-0.005		
		(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
Ethnic and Education (wife) FE	Yes	Yes	Yes	Yes

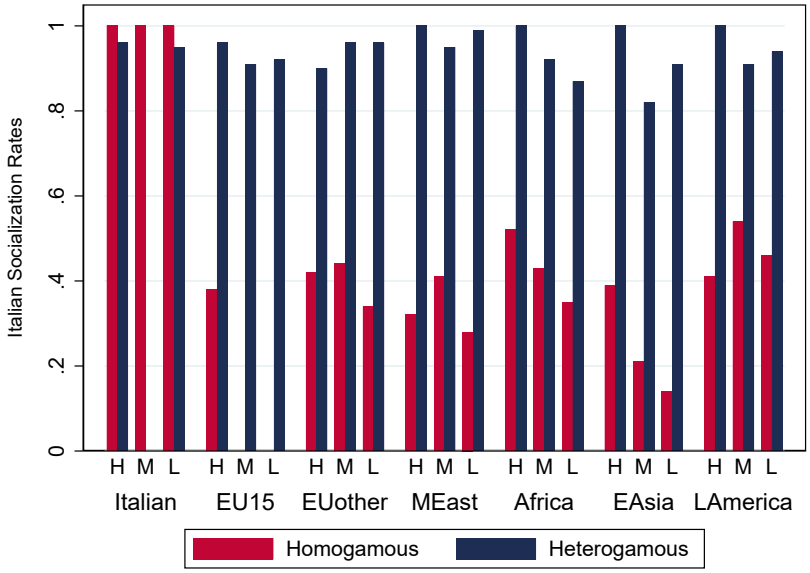
Notes: This Table shows the results of the regression of the fertility rate - in marriages defined by spouses' cultural ethnic-group and education, and by region - on various (combinations of) explanatory variables. In columns (1) and (2), the explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and education (*Homogamous by education*). Column (3) includes dummies for marriages by education, dividing marriages of high-educated (*High-high educ*) and mixed education marriages of (*High and low educ*) spouses. Column (4) includes interaction variables. All specifications include marriage market (at the regional level) fixed effects, wife educational level, and cultural-ethnic group of origin fixed effects. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.4: 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
Ethnic and Education (wife) FE	Yes	Yes	Yes	Yes	Yes

Notes: This Table shows the results of the regression of the separation rate - in marriages defined by spouses' cultural ethnic-group and education, and by region - on various (combinations of) explanatory variables. In columns (1)-(3) the explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and by education (*Homogamous by education*). In column (4), we include a dummy for marriages with at least one highly-educated spouse (*High educ, at least one*) and its interaction with the dummy for cultural homogamous marriages (*Hom by culture × High educ, at least one*) in column (5). 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.5: Italian Socialization Probabilities by Ethnic Group and Education of Spouses



Notes: This Figure shows socialization probabilities to Italian culture by ethnic group and education of spouses (H: both highly-educated spouses, M: high and low-educated spouses, L: both spouses with low education). The outcome variable is "Italian spoken at home." Estimates are reported separately for homogamous (red) and heterogamous (blue) households.

Table C.5: 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
Ethnic and Education (wife) FE	Yes	Yes	Yes	Yes

Notes: This Table shows the results of the regression of "Italian spoken at home" - in marriages defined by defined by spouses' cultural ethnic-group and education, and by region - on various (combinations of) explanatory variables. In columns (1) and (2), the explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and education (*Homogamous by education*). Column (3) includes dummies for marriages by education, dividing marriages of high-educated (*High-high educ*) and mixed education marriages of (*High and low educ*) spouses. Column (4) includes interaction variables. All specifications include marriage market (at the regional level) fixed effects, wife educational level, and cultural-ethnic group of origin fixed effects. 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: 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

Notes: This Table shows the results of the regression of "highly-educated children" - in marriages defined by defined by spouses' cultural ethnic-group and education, and by region - on various (combinations of) explanatory variables. In columns (1)-(3), the explanatory variables are dummies for homogamous couples by culture (*Homogamous by culture*) and by education (*Homogamous by education*). In column (4), we includes a dummy for marriages with at least one highly educated spouse (*High educ, at least one*) and its interaction with the dummy for cultural homogamous marriages (*Hom by culture × High educ, at least one*) in column (5). All specifications include marriage market (at the regional level) fixed effects, wife educational level, and spouses' cultural-ethnic group of origin fixed effects. Standard errors clustered at the regional level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.7: Italian Socialization Probabilities by Ethnic Group and Marital Status

	Italian Socialization Probabilities			
	Homogamous Families		Heterogamous Families	
	Married	Separated	Married	Separated
Italian	1	1	0.936	0.736
Europe-EU15	0.410	0.546	0.885	0.750
Other Europe	0.389	0.472	0.940	0.786
North Africa-Middle East	0.268	0.357	0.919	0.619
Sub-Saharan Africa	0.398	0.238	0.927	0.600
East Asia	0.198	0.242	0.856	0.375
Latin America	0.493	0.426	0.927	0.750

Notes: This table shows Italian socialization probabilities by ethnic group of spouses and marital status. The outcome variable is an indicator for whether the child speaks Italian within the family. Estimates are reported separately for homogamous and heterogamous married and separated families.

Table C.8: 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.

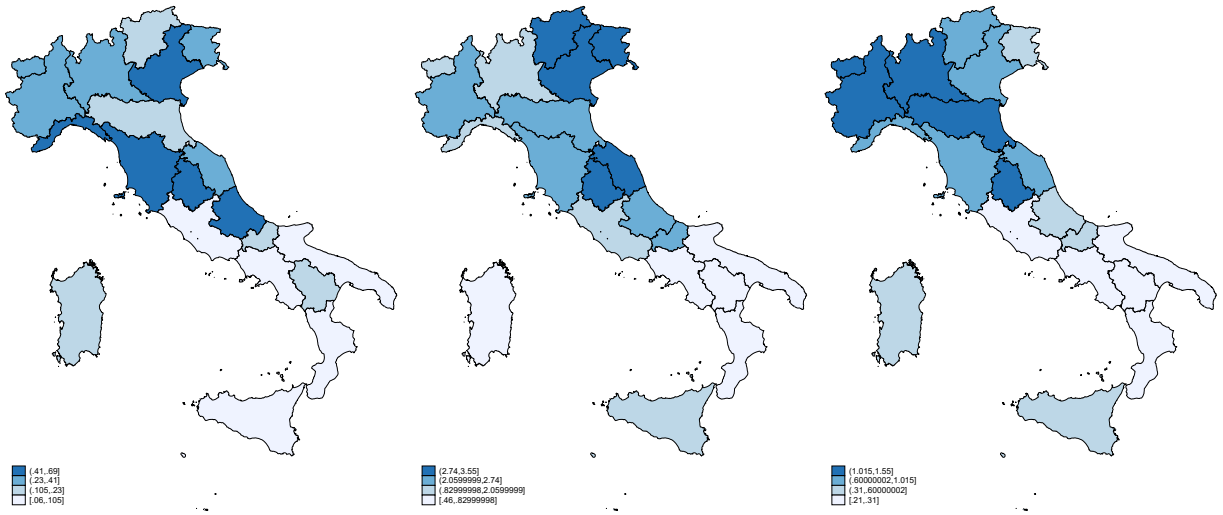
Table C.9: Singles by Cultural-Ethnic Group and Education

	Single Males				Single Females			
	Unskilled (%)		Skilled (%)		Unskilled (%)		Skilled (%)	
Italian	1994972	51,73	3366578	60,87	996934	40,60	2898654	54,93
Europe-EU15	24963	55,59	31161	54,11	15739	42,89	33036	48,96
Other Europe	79040	69,63	85751	70,46	56212	40,45	111006	49,70
Middle-East	37887	69,43	22818	63,24	12098	57,58	10923	59,60
Sub-Saharan Africa	18024	70,31	14029	61,50	11051	47,57	9759	51,41
East Asia	35551	79,94	20331	78,82	17549	58,06	16259	55,09
Latin America	17128	58,70	21777	57,24	20356	36,65	32726	41,98

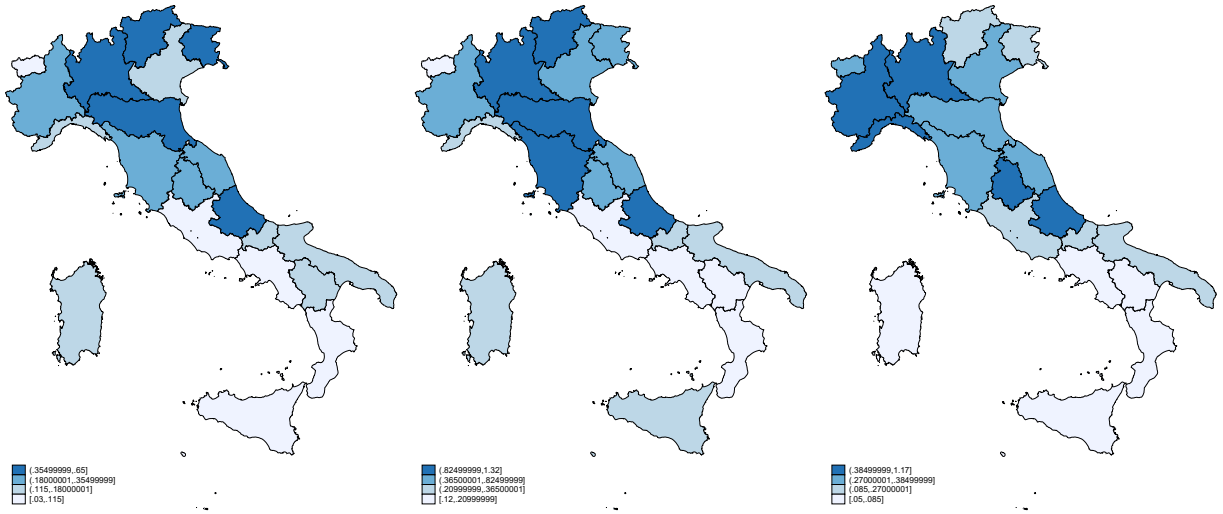
Notes: This Table reports the distribution of singles by gender and cultural-ethnic group, separately for individuals with low and high education. Singles are defined as adult individuals not legally married over the 90th percentile of the age at marriage distribution. We report absolute numbers and shares, computed as the number of singles over the total number of individuals by gender, cultural-ethnic group, and education.

Figure C.6: Migrants' Distribution across Regions

(a) Europe - EU15 (b) Other Europe (c) North Africa-Middle East



(d) Sub-Saharan Africa (e) East Asia (f) Latin America

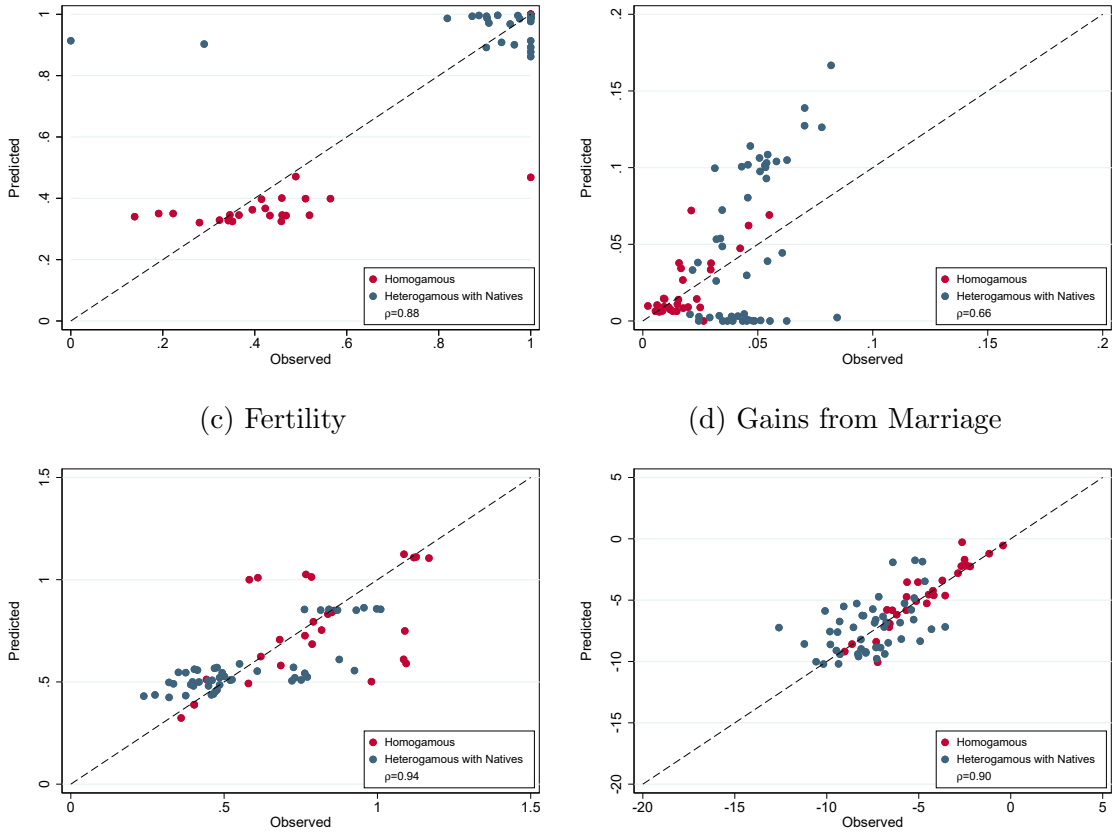


Notes: This Figure shows the distribution of the immigrant population by cultural-ethnic group and region. Population shares by cultural-ethnic group and region are computed over the total resident population at the regional level. The ethnic group classification is defined in Table C.1. The color classification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

Figure C.7: Fit of the model

(a) Socialization

(b) Separations

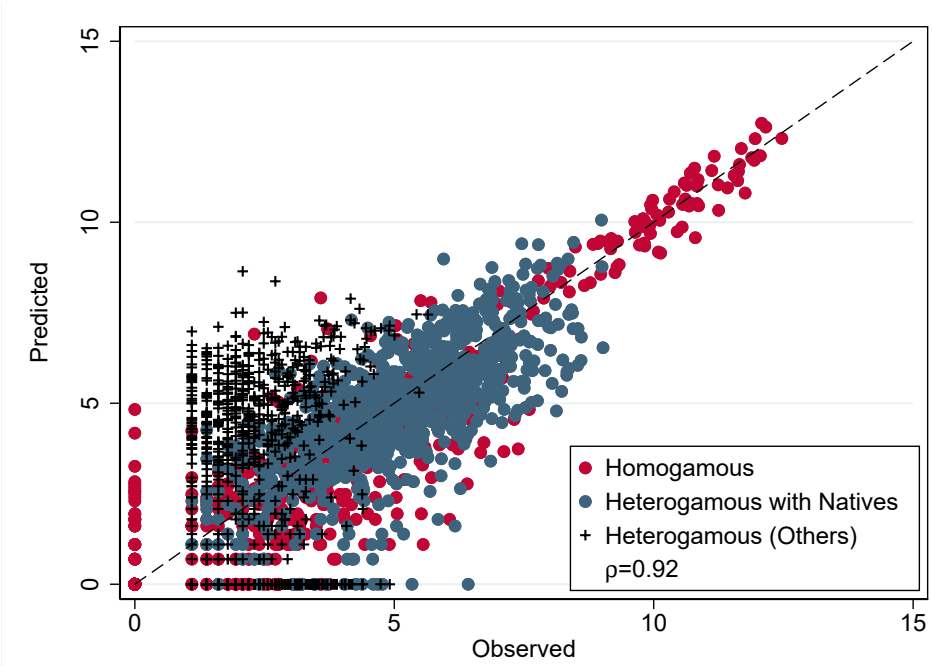


(c) Fertility

(d) Gains from Marriage

Notes: This Figure shows the average fit of the model by hj household, considering socialization rates to native Italian 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).

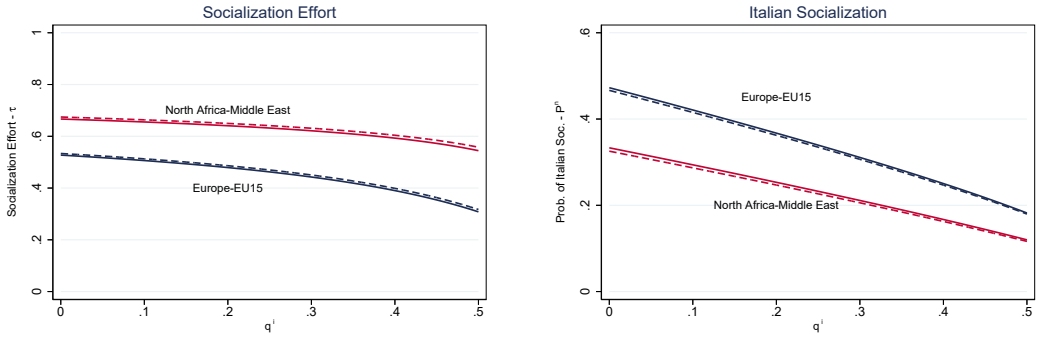
Figure C.8: Fit of the Model - Number of Marriages by Match and Region



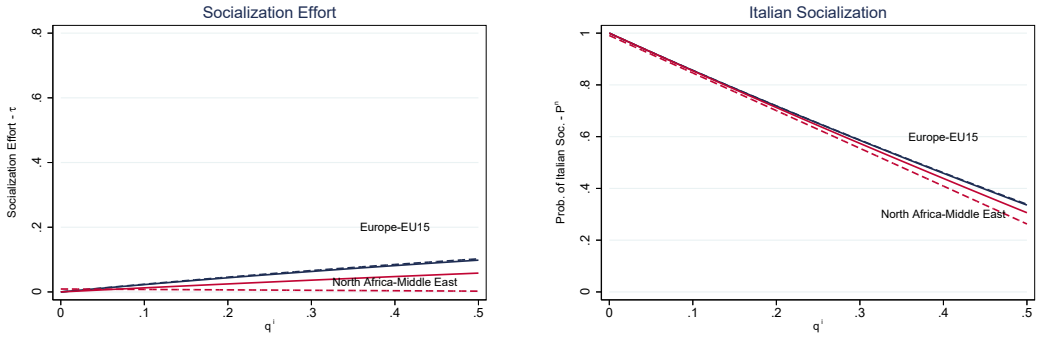
Notes: This Figure shows the relationship between the number of marriages observed in the data (in log) and the number of marriages predicted by the model (in log) by region for homogamous families (red), heterogamous families with natives (blue), and all other heterogamous matches (black).

Figure C.9: 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 socialization probability to Italian culture, P^n, over the potential population share, q^i, for European-EU15 and North African-Middle East minorities. Results for homogamous and heterogamous marriages with natives are, respectively, in panel a) and b).

Table C.10: Outside Option of Singlehood Parameters

Cultural group c:	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah Africa	East Asia	Latin America
				Homogamous households			
Outside option, High-High Educated ω_c	75.97	59.08	14.48	24.63	36.78	13.33	12.047
	(1.893)	(2.097)	(1.097)	(2.207)	(1.782)	(1.972)	(1.250)
Outside option, Mix educated ω_c	74.85	63.45	22.87	30.69	38.13	20.38	8.53
	(1.859)	(2.431)	(1.446)	(2.374)	(2.042)	(2.128)	(1.664)
Outside option, Low-Low educated ω_c	71.86	63.28	26.36	33.32	35.01	25.57	5.17
	(1.785)	(2.282)	(1.152)	(1.974)	(1.688)	(1.430)	(1.401)
Economic complementarity in marriage α_{hom}	0.425						
	(0.017)						
				Heterogamous households			
Outside option for het, ω_c	47.12	85.48	7.51	0.90	14.02	5.36	7.76
	(0.029)	(0.029)	(0.030)	(0.029)	(0.029)	(0.029)	(0.028)
Outside option ω_e High-High Educated	0.696						
	(0.018)						
Outside option ω_e Mix Educated	0.682						
	(0.017)						
Outside option ω_e Low-Low Educated	0.629						
	(0.015)						
Economic complementarity in marriage α_{het}	0.427						
	(0.039)						

Notes: This Table shows structural parameter estimates of the outside option of being single ω_{cz} and $\bar{\omega}_{cz}$. Standard errors are reported in parentheses.

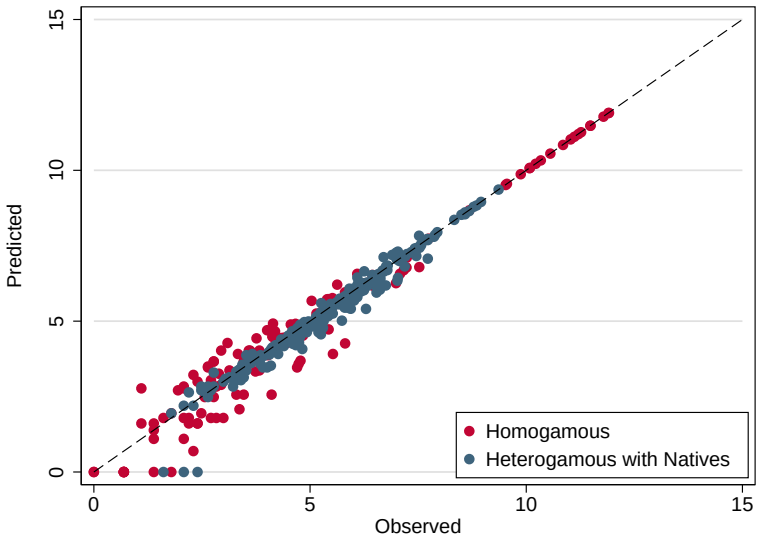
Table C.11: Structural Model Parameters - Model of Marriage by Culture Only

Cultural Intolerance Parameters							
c:	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah Africa	East Asia	Latin America
ΔV_n^c, Italian		33.38	58.60	67.88	78.23	61.53	19.27
$\Delta V_{i_E}^c$, Europe-EU15	10.21		52.60	4.77	6.69	18.54	0.33
$\Delta V_{i_O}^c$, Other Europe	39.97	0.05		69.32	56.63	30.42	23.02
$\Delta V_{i_M}^c$, North Africa-Middle East	65.35	7.00	58.98		97.85	52.12	52.25
$\Delta V_{i_A}^c$, Sub-Saharan Africa	55.00	28.37	58.42	96.37		81.22	42.50
$\Delta V_{i_S}^c$, East Asia	40.02	0.30	87.08	54.70	47.29		93.13
$\Delta V_{i_L}^c$, Latin America	38.95	10.69	20.77	58.99	46.96	29.13	

Cost Function and Extra Parameters					
Socialization Cost Parameters	σ_τ hom	10.82	Fertility Cost Parameters	σ_n hom	67.62
	λ_τ hom	0.549		λ_n hom	0.006
	σ_τ het	21.61		ϵ hom	1.021
	λ_τ het	0.571		σ_n het	99.86
Extra Marital Gain per Child	δ	0.820		λ_n het	0.021
Segregation Parameter	ρ	1.765		ϵ het	1.229
Economic complementarity in marriage	α	0.484			

Notes: This Table shows structural parameter estimates from a model of marriage along cultural-ethnic lines only; that is, with $\gamma_c = 1$ and $S_c = 0$ for all c.

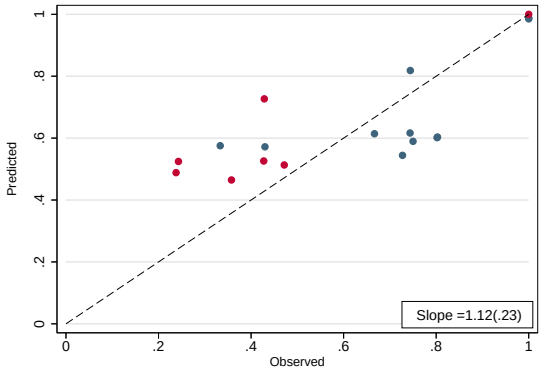
Figure C.10: Model Validation - Non-Targeted 2013-2019 Marriages by Match and Region



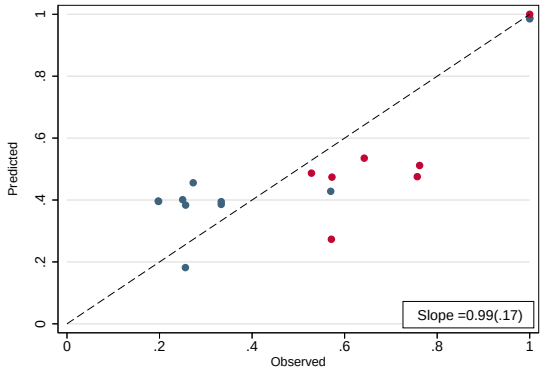
Notes: This Figure displays the scatterplot of the relationship between the number of marriages observed in the out-of-sample data (in log) in the years from 2013 to 2019 and the number of marriages predicted by the model (in log) by region, for homogamous families (red) and heterogamous families with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

Figure C.11: Model Validation - Non Targeted Socialization Rates for Divorced Couples

(a) Italian Socialization



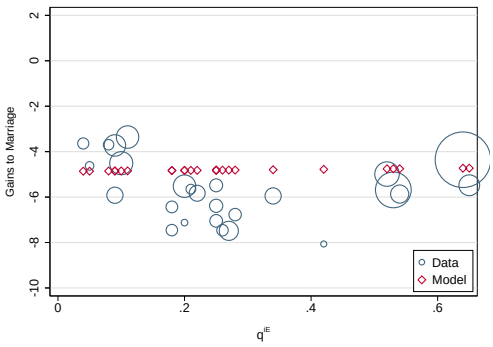
(b) Mother Socialization



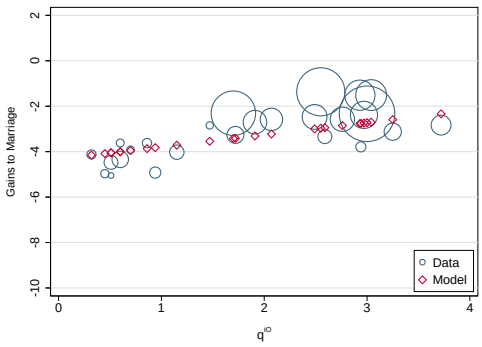
Notes: This Figure displays the scatterplot of the relationship between the observed and predicted Italian and mother socialization probabilities for the subsample of marriages ending in divorce, separately for homogamous families (red), and heterogamous families with natives (blue); 45-degree dashed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

Figure C.12: Model Validation - Gains from Marriage for Homogeneous Families by Province

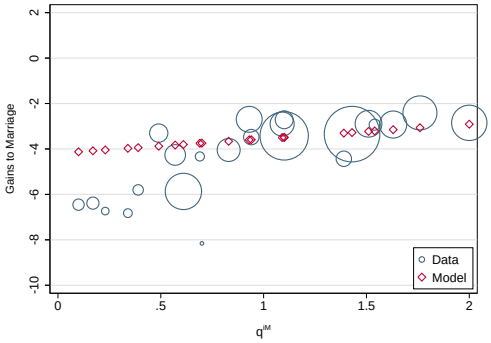
(a) Europe-EU15



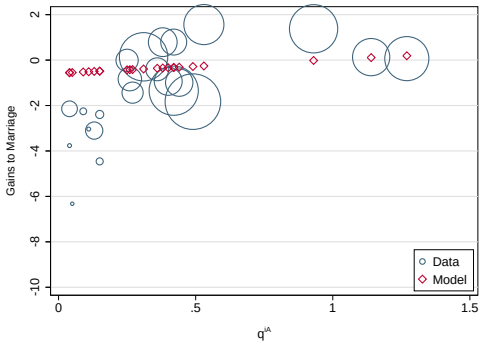
(b) Other Europe



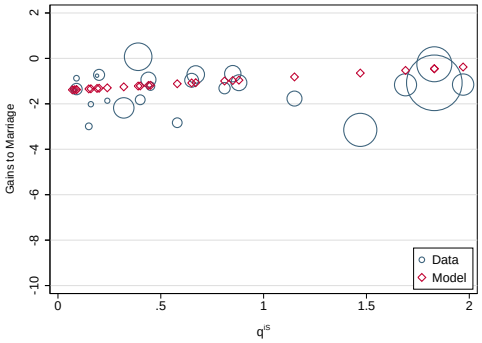
(c) North Africa-Middle East



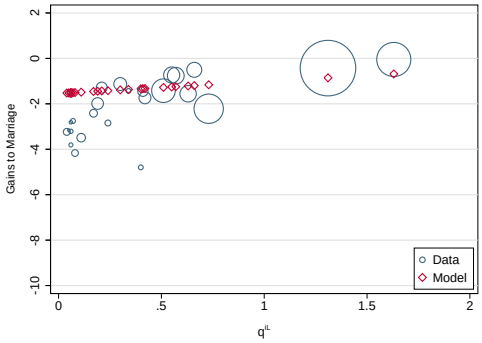
(d) Sub-Saharan Africa



(e) East Asia



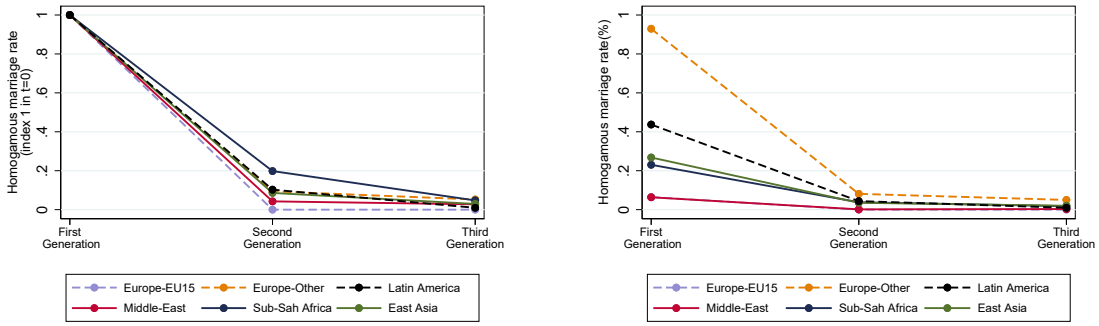
(f) Latin America



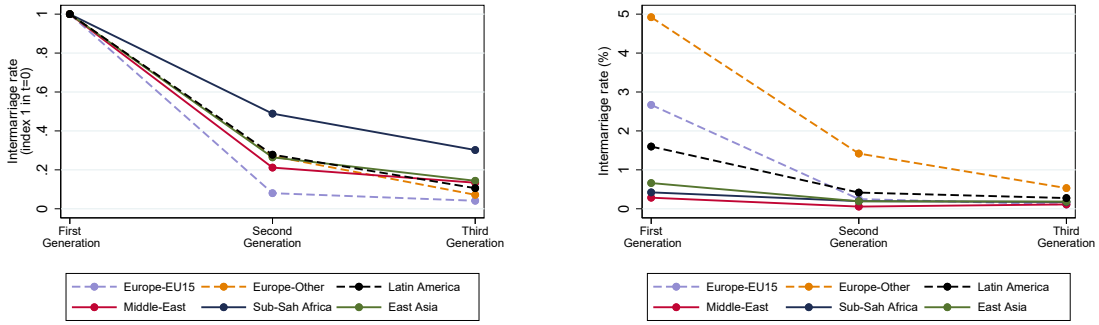
Notes: The Figure shows out-of-sample predicted and implied gains from marriage for homogamous families of ethnic group minorities over the corresponding population share, q^i (in percentage), by province of residence (average over the time period). Empirical moments are weighted by the observed number of marriages per province. We select the most representative provinces across the 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.13: Inter-generational of Marital Matching

(a) Homogamous Marriages

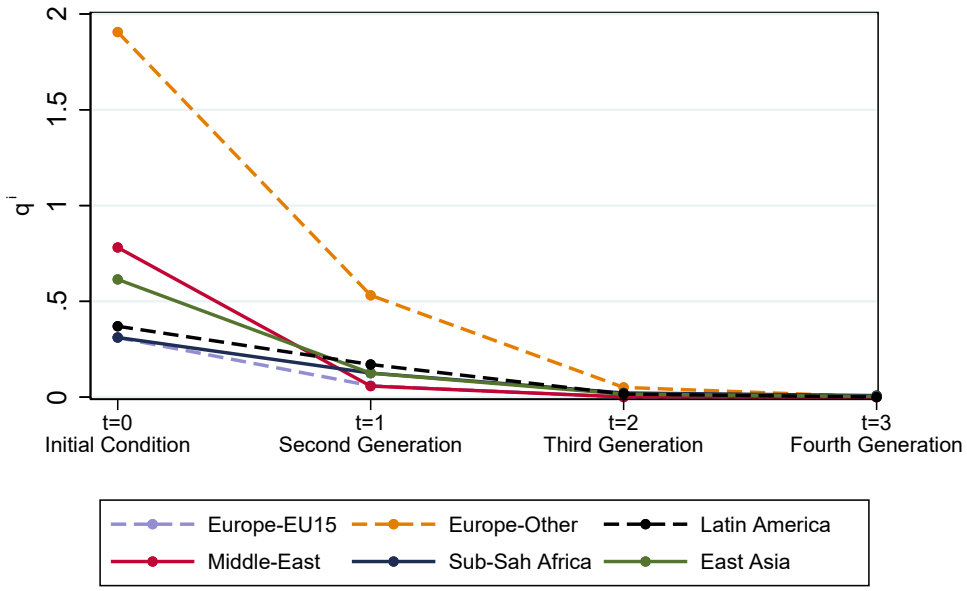


(b) Heterogamous Marriages



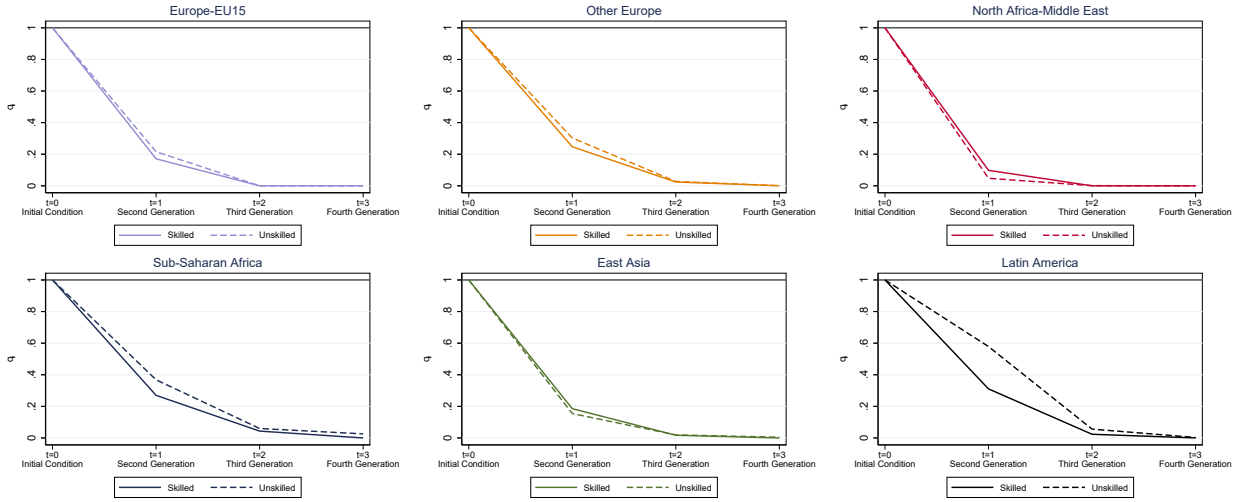
Notes: This Figure shows the inter-generational dynamics of matching patterns for homogamous marriages and heterogamous marriages with natives, in panel a) and b), respectively, over successive generations.

Figure C.14: Inter-generational Dynamics of Cultural Traits



Notes: This Figure shows the inter-generational dynamics of the distribution of cultural traits in the population for all minorities.

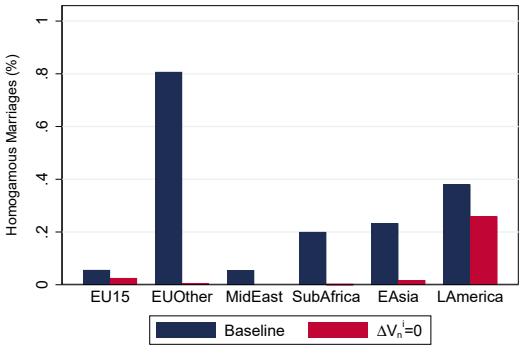
Figure C.15: Inter-generational Dynamics of Cultural Traits by Education



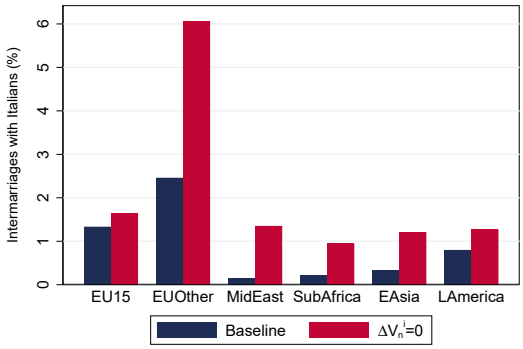
Notes: This Figure shows the inter-generational dynamics of the distribution of cultural traits in the population for minority groups by education, normalized so that $q_t^{ie} = 1$ in $t = 0$ for $e \in \{s, u\}$. The solid line represents the dynamics for highly-educated immigrant minorities, while the dashed line represents the dynamics for immigrant minorities with low education.

Figure C.16: Change in Matching Patterns with Italians Fully Tolerant

(a) Homogamous marriages



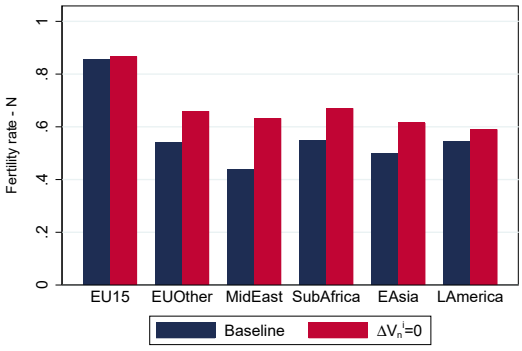
(b) Intermarriages with Natives



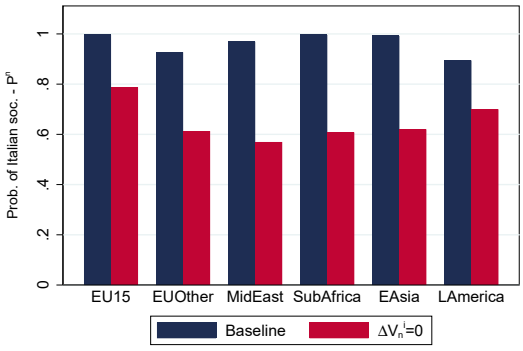
Notes: This Figure shows the percentage change in homogamous and heterogamous marriages, in panel a) and b), respectively - with full tolerance of natives towards minorities with respect to baseline; that is, $\Delta V_{ns}^{is} = 0$.

Figure C.17: Change in Intra-household Patterns with Italians Fully Tolerant

(a) Fertility

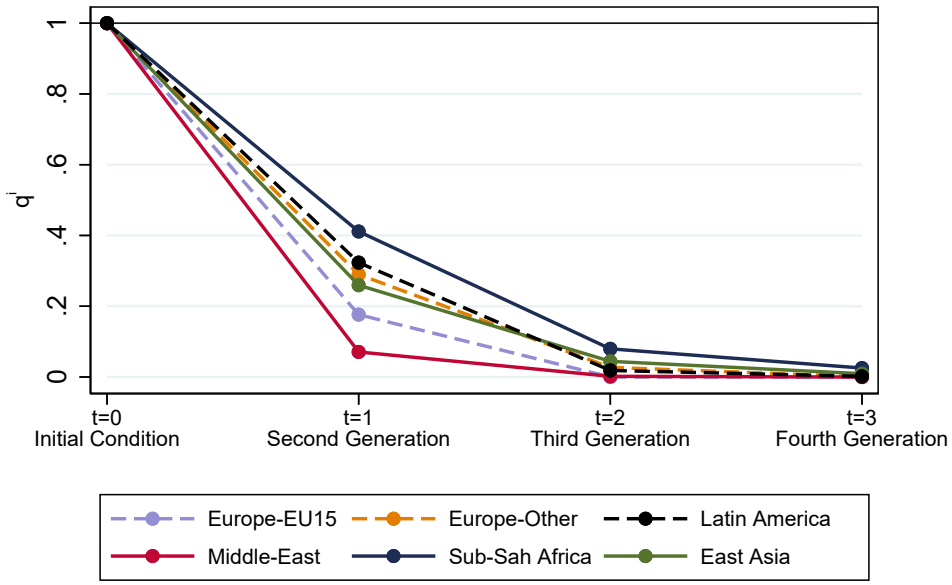


(b) Italian socialization



Notes: This Figure shows the variation in fertility rate and socialization probability to Italian culture in panel a) and b), respectively - in intermarriages with natives at the baseline and in case of complete tolerance of Italian majority towards minorities; that is, $\Delta V_{ns}^{is} = 0$.

Figure C.18: Inter-generational Dynamics of Cultural Traits with Italians Fully Intolerant, $\Delta V_{ns}^{is} = 100$



Notes: This Figure shows the inter-generational dynamics of the distribution of cultural traits in the population for all minorities i, assuming the case of complete intolerance of Italian majority towards minorities.

Table C.12: 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 linguistic socialization to Italian culture ("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, gender and age fixed effects. Standard errors clustered at the province level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.13: 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 linguistic socialization to Italian ("Italian at home") influences the achievement and educational choices of immigrant students. The dependent variables include the reading and math standardized test scores in grade 5 in panel a) and b), respectively, and a dummy equal to one for students choosing the high-track (academic or technical schools) 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. Estimates are weighted by a cheating correction factor, except in column (4). 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.

D Additional Robustness

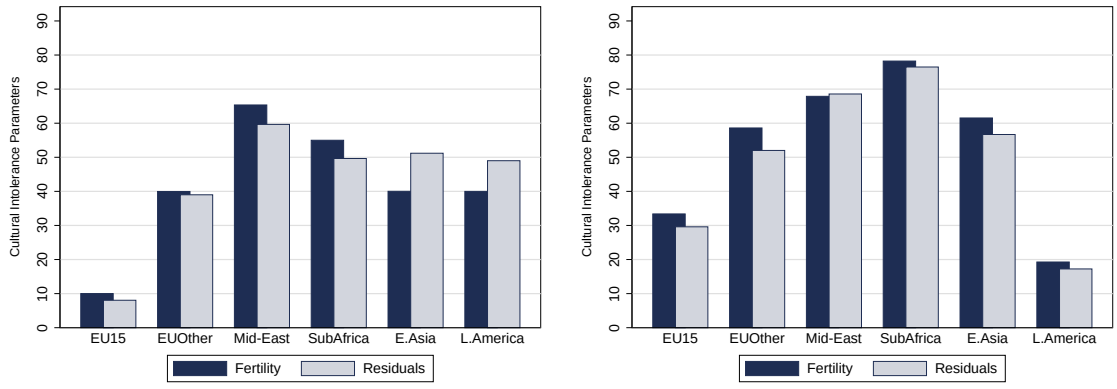
We present three robustness exercises to address the possibility that: (i) fertility differences may reflect systematic variation in spouses' observable characteristics; (ii) the population distribution across regions may be endogenous; and (iii) children's educational outcomes may depend on their parents' marital status, that is, whether the parents are divorced.

First, we address the possibility that differences in fertility across cultural and education traits might be due to systematic variation in observables across households.[60] We regress fertility on covariates such as marital duration, age at marriage, educational attainment (primary education, secondary education, high-school education level, college degree and above), a dummy variable for employment, and occupational standing (blue-collar, white-collar, director, self-employed, and entrepreneur). We compute fertility moments from regression residuals, which control for confounding factors. Figure D.1 shows cultural intolerances consistent with our main estimates, with the exception of a slight underestimation of the demand for cultural identity among East Asian minorities.

Figure D.1: Cultural Intolerance Parameters - Robustness with Fertility Residuals

(a) Migrants towards Natives (Demand parameters)

(b) Natives towards Migrants (Supply parameters)



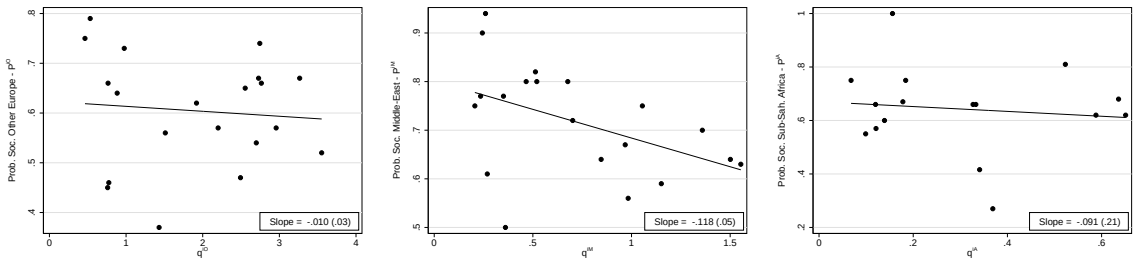
Notes: This Figure reports parameter estimates of the cultural intolerance of immigrants versus natives ΔV_i^n and natives versus immigrants ΔV_n^i - in panel a) and b), respectively - for all cultural-ethnic groups i. The blue bars report baseline estimates. The grey bars, instead, report estimates 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 (primary education, secondary education, high-school education level, college degree and above), a dummy variable for employment, and occupational standing (blue-collar, white-collar, director, self-employed, and entrepreneur).

[60] As in the validation exercises in Section 5.3, we disregard educational variability. Theoretically, we impose $\gamma_c = 1$ and $S_c = 0$, and re-estimate the other parameters in our model; see footnote 39.

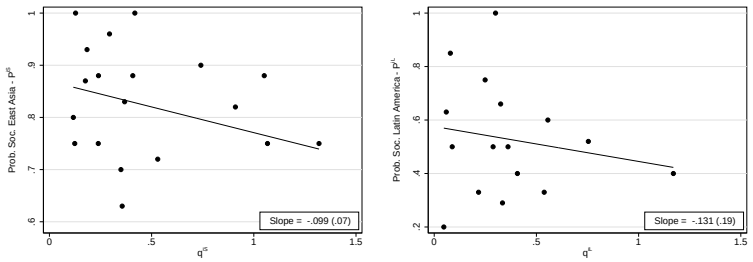
Our second robustness exercise tackles the possible endogeneity of the population distribution across regions. Endogenous moving or location decisions would be problematic for our estimates if these decisions were partly motivated by marriage, socialization, and unobserved heterogeneity. Consider the 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 direct vertical and indirect horizontal socialization. Figure D.2 describes this relationship, showing instead a negative and, at times, statistically significant correlation. To further mitigate this concern, we adopt a shift-share approach (Card, 2001; Tabellini, 2020), predicting ethnic group distributions across regions based pre-existing variation in immigrants' settlements by ethnic group observed in 1993, interacted with subsequent inflows by origin. Given the persistent ethnic clustering of immigrants, these historical patterns strongly predict location choices of newly arriving immigrants. The resulting estimates remain consistent with our baseline findings.

Figure D.2: 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.

In our final robustness analysis we investigate whether educational attainment may be

affected by the divorce decision and explore its implications for our model. We treat children's educational attainment as exogenous, directly observed in the data. In this revision, we maintain the same approach in modeling children's education conditional on divorce. Crucially, we rely on two data sources to estimate the probability that children have low education (i.e., enroll in low-track vocational schools as opposed to high-track schools), conditionally on parents' cultural-ethnic traits and educational background, by region: INVALSI data and ISTAT data from the Condition and Social Integration of Foreign Nationals Survey; see Appendix B for details on the INVALSI data. Unfortunately, the INVALSI data do not provide information on parents' marital status. However, the ISTAT survey does include this information - though for a relatively small sample of approximately 1,900 students between 15 and 25 years old (the target population for our analysis). With these data we estimate the probability of children having low education conditionally on parental divorce.

Results in column (1) of Table D.1 show that children with divorced parents are, on average, more likely to have low education, i.e., they are 5 percent more likely to attend low-track (vocational) schools. The effect is positive and statistically significant at the 10% level. The estimates control for cultural-ethnic fixed effects as well as region (marriage market) fixed effects. We further investigate heterogeneity by parental education. Specifically, we report estimates separately for households in which both spouses are highly educated (column 2), the husband is highly educated and the wife has low education (column 3), and the husband has low education and the wife is highly educated (column 4). Unfortunately, the proportion of divorced couples in households with spouses with low education is insufficient to support any meaningful estimation. Crucially, the effect of interest is driven solely by families with mixed education levels, and only when the mother has low education. We find no significant effects for the other household types.

Thus, while in principle, educational attainment may be affected by the parental divorce decision, the evidence appears to be relevant only for a specific subset of households, namely, those with mothers with low education, who are more likely to retain custody of the children after divorce and raise them in a context with lower educational resources. This reasoning does not appear to generalize across all households in our data. However, we again emphasize that information on divorce is available only from the ISTAT survey, which covers a relatively small sample of divorced families. As a result, the estimated probabilities of educational attainment conditionally on divorce are subject to measurement error and should be interpreted with caution.

Despite these data limitations, we have re-estimated the model, allowing for different probabilities of educational attainment conditionally on divorce status. This approach accounts for the empirical variation in educational outcomes associated with parental marital

status. We report our estimates in Table D.2. We highlight three points. First, the estimate of δ—the extra value of each child attributable to staying in a marriage, as opposed to divorcing, independently of socialization—more than doubles, from 0.82 to 1.98. This is to be expected, as this parameter now captures the fact that staying married leads to an education premium, i.e., it increases the probability that children attain a high level of education, which has value for parents. Second, cultural intolerance parameters are consistent with our main estimates, see Figure D.3. Finally, the educational parameters S_c also remain in line with our main estimates, and the γ_c parameters are estimated to be close to or even greater than one. That is, cultural intolerance tends to be equal to or higher among immigrant parents with low education compared to those with high education.

Table D.1: Estimates of Children with Low Education by Spouses' Education and Marital Status

	(1) All sample	(2) High-high	(3) High and Low (Low educ mother)	(4) Low and High (High educ mother)
Dep var:	Probability of "Low educated" Children			
Divorce (dummy)	0.051*	-0.039	0.091*	-0.019
	(0.027)	(0.052)	(0.047)	(0.067)
	{0.061}	{0.454}	{0.053}	{0.778}
Observations	1916	408	380	309
R-squared	0.062	0.129	0.060	0.087
Marriage market (region) FE	Yes	Yes	Yes	Yes
Ethnic FE	Yes	Yes	Yes	Yes

Notes: This Table reports OLS estimates of the the probability of low educated children by marital status of the parents. The dependent variable is a dummy equal to one for low educated children, i.e., students attending low-track (vocational) schools. The sample is restricted to students between 15 and 25 years old. The main explanatory variable is a dummy equal to one for separated parents. Column (1) reports average estimates. Columns (2)-(4) report estimates separately for households by spouses' education, distinguishing households with both highly-educated spouses (column 2), households with highly-educated husband and low-educated wife (column 3) and households with low-educated husband and highly-educated wife (column 4). All regressions include marriage market (at the regional level) fixed effects and ethic-group fixed effects of both spouses. Robust standard errors clustered at the regional level are reported in parentheses; p-values are reported in curly brackets. Significance level: * p<0.1.

Table D.2: Structural Model Parameters - Robustness with Education conditional on Divorce

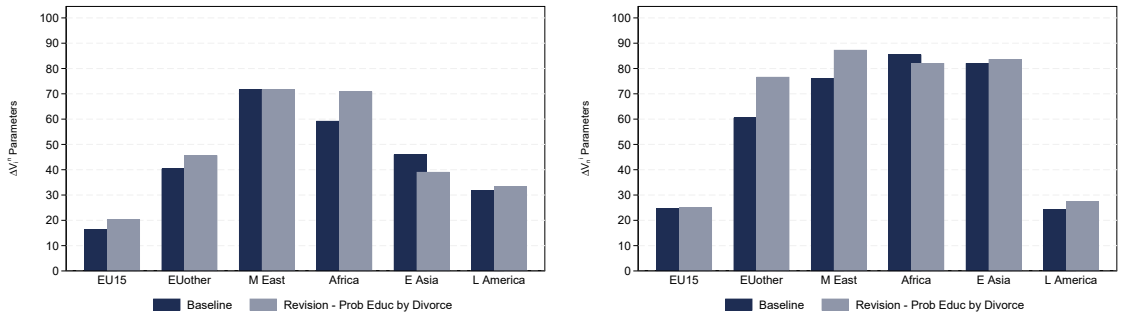
			Cultural Intolerance Parameters				
c:	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah Africa	East Asia	Latin America
ΔV_n^c, Italian		24.96	76.64	87.31	82.10	83.65	27.49
ΔV_{iE}^c, Europe-EU15	20.45		48.51	14.76	87.31	87.02	17.43
ΔV_{iO}^c, Other Europe	45.62	25.73		95.66	88.14	89.64	68.82
ΔV_{iM}^c, North Africa-Middle East	71.56	9.31	97.73		54.55	59.90	79.29
ΔV_{iA}^c, Sub-Saharan Africa	70.83	95.56	80.94	46.11		59.73	59.08
ΔV_{iS}^c, East Asia	39.08	46.28	85.83	63.74	29.65		56.42
ΔV_{iL}^c, Latin America	33.20	15.48	67.52	90.64	76.62	62.93	
			Education Parameters				
Surplus for skilled children, S_c	15.29	10.67	45.09	11.08	1.435	34.97	45.61
Educational preference of unskilled parents, γ_c	1.037	1.037	1.600	1.088	0.973	1.359	0.936
			Cost Function and Extra Parameters				
Socialization Cost Parameters	σ_τ hom	12.17	Fertility Cost Parameters			σ_N hom	67.27
	λ_τ hom	0.921				λ_N hom	0.099
	σ_τ het	23.46				ϵ hom	1.027
	λ_τ het	0.433				σ_N het	82.19
Extra Marital Gain per Child	δ	1.988				λ_N het	0.199
Segregation Parameter	ρ	1.494				ϵ het	1.017

Notes: This Table shows the estimates of the structural parameters of the model.

Figure D.3: 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} and natives versus immigrants ΔV_{ns}^{is}, in panel a) and b), for natives n and all cultural minorities i. The blue bars report baseline estimates. The grey bars, instead, report estimates allowing in the model for different probabilities of educational attainment conditionally on divorce status.

E Identification

We provide here 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 traits c, natives and immigrants $c \in (n, i)$. 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). The following detailed analysis of identification is summarized in Table E.1.

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.[61] 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 of symmetric):

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

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

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

Table E.1: Recursive Identification Structure of the Model

Parameter	**Theoretical Expression**
Socialization in homogamous skilled immigrant households	
$\frac{\Delta V_{is}^{ns}}{\sigma_\tau}$	$P_{isis}^{i} = \frac{2}{\sigma_\tau}\Delta V_{is}^{ns}(1-q_i) + \left(1 - \frac{2}{\sigma_\tau}\Delta V_{is}^{ns}(1-q_i)\right) q_i$
Socialization in mixed-education homogamous immigrant households	
γ_i	$P_{iuis}^{i}, P_{isiu}^{i}, P_{iuiu}^{i}$ depend on γ_i and ΔV_{is}^{ns}
Socialization in immigrant-native heterogeneous couples	
ΔV_{ns}^{is}	Based on $P_{isns}^{i} > q_i$, identify which parent socializes
Socialization in native-immigrant heterogeneous households	
γ_n	Linear in $\frac{\gamma_n \Delta V_{ns}^{is} - \Delta V_{is}^{ns}}{\sigma_\tau}$
Divorce in homogamous native households	
δ	$\pi_{hj}(N) = F(Nu_{hj}(\tau(1),1) - Nu_{hj}(\tau(0),0) - N\delta)$
Divorce and socialization in homogamous immigrant households	
σ_τ, S_i	Use known $u_{hj}(\tau, d)$, solve recursively for S_i, σ_τ
Fertility across all household types	
σ_N	$N_{hj} = \frac{1}{\sigma_N}[\pi u_{hj}(\tau(1),1) + (1-\pi)(\delta + u_{hj}(\tau(0),0))]$
Divorce probability without children across all household types	
a_{hj}	$F(0; a_{hj}, 1) = \hat{\pi}_{hj}(0)$
Gains from marriage across all household types	
$\omega_{hz}, \omega_{jz}, \alpha_z$	$G_{hj} = U_{hj} - \omega_{hz} - \omega_{jz}$ and $U_{hj}^{ec} = \alpha_z(\omega_{hz} + \omega_{jz})$

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 $d = 0$.

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

$P^i_{isns} > q^i$ 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.[62]

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

[62] 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 from homogamous households of natives. Consider now equations (Div) for homogamous households of immigrants, with $h = is;\ j = is$ and with $h = iu;\ j = is$, where

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

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

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 $h = is; j = is$. 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:

$$G_{hj} = \log \frac{(\mu_{hj})^2}{\mu_{h0} \cdot \mu_{0j}}; \tag{8}$$

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 μ_{h0}, μ_{0j} are data. These moments then depend on the parameters already identified and linearly on ω_{h_z} and ω_{j_z}. In this particular case, with 4 possible individual types $h, j = ce$ and two household types $z \in \{hom, het\}$, we have $4 \times 2 = 8$ outside options ω_{ce_z} and $4^2 = 16$ gains from marriage, so that the outside options are identified.

In the more general model, with 7 cultural-ethnic traits in the population, we also identify the relative effects of marriage on spouses' economic opportunities, α_z, with $z \in \{hom, het\}$. Given our specification,

$$U_{hj}^{ec} = \alpha_z \left(\omega_{h_z} + \omega_{j_z}\right);$$

for each household type hj, the moment equations for the implied marital surplus net of outside options are:

$$U_{hj} + (\alpha_z - 1)(\omega_{h_z} + \omega_{j_z}) = G_{hj}.$$

In this system, with 14 possible individual types $h, j = ce$ and two household types z, we have $14 \times 2 = 28$ outside options plus 2 α_z parameters and $14^2 = 196$ estimated gains from marriage, so that the outside options and the α_z parameters are overidentified.