Marriage, Fertility, and Cultural Integration in Italy*

Alberto Bisin† Giulia Tura‡

March 15, 2022

Abstract

We study cultural integration as an equilibrium of marital matching and intrahousehold decisions regarding fertility and cultural socialization. Structural estimates reveal strong demand to preserve cultural identity on the part of immigrants and little acceptance of the immigrants' cultural diversity of natives. Nonetheless, we simulate substantial, though heterogeneous, integration rates across immigrant groups, 75% on average over one generation. Counterfactuals show how more accepting preferences of the natives would lead to slower cultural integration, while a reduction in economic incentives to immigrants would increase it. Finally, we evaluate a policy enhancing social welfare by strengthening the ethnic network of immigrants.

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

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

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

1 Introduction

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

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

To this end, we are able to exploit rich administrative marriage data providing information on the cultural-ethnic identity of immigrants in Italy. These micro-level data, provided by the Italian Statistical Institute (ISTAT, ADELE Laboratory), cover the universe of marriages formed in Italy from 1995 to 2012 and the universe of births and separations registered in the same period.[3] In addition, we recover a measure of parental cultural socialization of children from the *Condition and Social Integration of Foreign Nationals* Survey (2011-2012). Specifically, we interpret "speaking Italian at home" as a proxy for (negative) socialization rates to the culture of origin, that is, an indication that parents are not exercising much effort in transmitting their culture to their children.[4]

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

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

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

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

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

riages displaying systematically higher childbirth investments. Moreover, these patterns are negatively associated to the cultural distance of the ethnic traits of the spouses.

To separately identify the cultural preferences of immigrants and natives, we then study a structural model of the marriage market in which spouses match along the cultural-ethnic dimension and solve a collective household decision problem determining fertility and children socialization. Spouses match along cultural-ethnic lines, identified by individuals' culture of origin. Marital utility depends on observable spouses' characteristics and results from a collective household decision problem embedded within this matching framework. Within marriage, parents choose fertility, investments in the cultural socialization of children, and possibly divorce. Parents who value their cultural identity care about socializing their children, and they are endowed with socialization technologies to transmit their own cultural-ethnic traits to children. Thus, parents choose to exert a direct socialization effort to affect their children's process of cultural identity formation. But effort requires parental resources, e.g., time spent with children, private school tuition, selection of residential neighborhoods and ethnic network and so on. Socialization incentives and technologies vary, in particular, between homogamous and heterogamous marriages. For instance, families where parents share the same cultural traits enjoy a more efficient socialization technology in transmitting their shared trait than families where parents do not share the same culture. Furthermore, parental socialization choices depend on the distribution of the population across ethnic groups. As a consequence, the model implies a systematic dependence of marriage, fertility, socialization, and divorce patterns, across households' cultural-ethnic characteristics.

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

Estimated cultural intolerance parameters are positive, asymmetric, and highly heterogeneous across cultural-ethnic groups. In other words, preferences for cultural identity dominate economic (and legal) incentives to integration for all minorities. This is particularly so for immigrants from North Africa-Middle East, whose estimated cultural intolerance is nearly seven times as high as Europeans'. On the other hand, we also estimate high cultural intolerances for the Italian majority; that is, little cultural acceptance overall. In particu-

lar, 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).

We investigate the evolution of the distribution of the population by cultural traits in the long-run, by simulating our model at the estimated parameters over successive generations. Despite high cultural intolerance estimates, all cultural-ethnic minorities are simulated to integrate to the Italian majority. The integration rate, defined as the reduction in the fraction of the total population (immigrants and natives) which is composed of immigrants who are not integrated to the native Italian culture, is 75% in one single generation. However, the pace of convergence is heterogeneous across cultural-ethnic groups. On the one hand, we find that minorities from Europe and North Africa-Middle East integrate rapidly. A slower integration rate characterizes instead immigrants from East Asia, Sub-Saharan Africa minorities, and especially Latin America, which reaches a 70% integration rate only after four generations. The patterns of cultural integration across ethnic groups in the simulations are not only the result of demand and supply effects at equilibrium; that is, of the cultural intolerance of immigrants towards the natives and of that of the natives towards the immigrants. Immigrants from North Africa-Middle East, Sub-Saharan Africa, and East Asia, for instance, have relatively comparable estimated cultural intolerance but significantly different dynamics of integration due to a substantial heterogeneity in their rates of homogamy and fertility. Indeed, our empirical analysis clearly shows that both selection into homogamous marriages and selection through fertility act as socialization mechanisms, affecting the dynamics of integration of cultural-ethnic minorities.

To examine in depth the mechanisms driving integration at equilibrium, we proceed by counterfactual analysis, connecting outcomes to reduced form results from the literature on immigrants' assimilation. More specifically, we study how integration responds to variations both in the supply of acceptance of the immigrants' cultural diversity on the part of natives, as well as in the demand of immigrants to preserve their cultural identity. Our first counterfactual simulation entails setting the cultural intolerance of natives with respect to all ethnic minorities equal to zero, so that natives offer complete acceptance of the immigrants' cultural diversity. In this case, we observe no integration of second-generation immigrants: the fraction of the population composed of non-integrated immigrants increases by 15% over the period of a generation. Facing no bias in the marriage market with natives, nor any preferences for culturally integrated children on the part of the native spouse in intermarriages, immigrants can achieve higher socialization rates when married with natives. At equilibrium, heterogamy increases, and so does socialization and fertility in intermarriages, reducing integration overall. In our second counterfactual simulation, we strengthen the dominance of cultural identity in the demand of immigrants, which could be due to a re-

duction in their economic incentives to integration. Interestingly, in this case, we find that cultural convergence is accelerated in a period of a generation, by 10 percentage points compared to baseline for a 20% increase in cultural intolerances. This result is the outcome of a lower participation in the marriage market and lower fertility of immigrants, motivated by the fact that the stronger attachment to their identities makes marriage a riskier and costlier investment. This is the case for both homogamous marriages of immigrants and heterogamous marriages. In other words, in this counterfactual, the acceleration of integration is mostly an effect of the reduction in the population growth for immigrants with respect to natives. The probability that a child with an immigrant parent is integrated to the Italian culture is lower in the counterfactual with higher cultural intolerance, because immigrant parents' socialization is more effective, but the fraction of the total population (immigrants and natives) belonging to the native culture, on the contrary, is higher.

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

Finally, we take advantage of our model of cultural integration to investigate the effects of a surge in migration inflows on cultural heterogeneity. Doubling the number of second-generation minorities (keeping population shares constant across minorities), the integration

rate lowers to 86% by the third generation compared to the 93% at the baseline. The effects are heterogeneous across cultural-ethnic groups: the integration rate is reduced by 20 and 6 percentage points, respectively, for immigrants from Sub-Saharan Africa and East Asia.

After a discussion of the related literature in the next section, the paper is organized as follows. Section 2 introduces an empirical analysis of marriage, fertility and cultural-ethnic socialization of children by cultural-ethnic group. Section 3 outlines our theoretical framework, and Section 4 presents the structural model, the estimation strategy, and the identification of model parameters. We present the estimation results and various counterfactual simulations in Section 5 and 6. The welfare analysis of segregation policies is in Section 7, and the simulations of migration inflows in Section 8. Finally, Section 9 concludes.

1.1 Related literature

Our paper combines insights from the literature on the intergenerational transmission of cultural traits with those of family economics studies on the estimation of marital preferences. Methodologically, we embed a collective household decision problem into a matching model, as first in Chiappori et al. (2017, 2018); other papers along these lines include Gayle and Shephard (2019) and Galichon et al. (2019).[5] Marital utilities emerge endogenously as a function of fertility and intra-household inter-generational socialization choices, along the lines of Bisin et al. (2004).[6] In this way, we account for the fact that fertility and child-rearing are two key motives behind marriage (Browning et al., 2014), and at the same time we investigate the mechanisms that make ethnic-cultural traits a crucial dimension of marital matching (Bisin et al., 2004; Ciscato and Weber, 2016).

In terms of research question, this paper fits into the large empirical literature on the cultural integration of immigrants. Several of these studies concentrate on the immigrants' demand to preserve their cultural identity, by exploring socialization via children's first names and home language (Abramitzky et al., 2020; Fouka, 2020), intermarriage patterns (Gordon, 1964; Meng and Gregory, 2005; Furtado and Trejo, 2013; Guirkinger et al., 2019), self-reported national identity (Manning and Roy, 2010), contraceptive usage by teenage females (Achard, 2020), and neighborhood sorting (Hwang, 2019). Relatedly, some papers

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

[6]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.

study the effects of specific immigration policies and reforms (Fouka, 2020; Abdelgadir and Fouka, 2019); others document the salience of the cultural identity across immigrant groups living in the same host country on a wide variety of outcomes (Fernández and Fogli, 2006, 2009; Fernández, 2011; Giuliano, 2007; Alesina et al., 2013). Separately, a sizable literature has also focused on the supply side of immigrants' integration, studying the economic roots of anti-immigrant sentiments on the part of natives (as surveyed in Borjas, 2014; Card and Peri, 2016; Dustmann et al., 2016), and their consequent political reactions (Dustmann et al., 2019; Tabellini, 2020). With respect to these studies, our structural approach allows us to identify and estimate both the demand of immigrants to preserve their cultural identity and the supply of acceptance of the immigrants' cultural diversity on the part of natives. Within this unified and coherent framework, we can then study how the dynamics of immigrants' integration over time respond in equilibrium to variations in the preferences of both immigrants and natives, and we can investigate the effects of counterfactual policies.

2 Marriage, fertility, and socialization by cultural group

This section introduces our empirical analysis of marriage, fertility, and socialization by cultural-ethnic group. We briefly present the data, referring to Appendix A for a comprehensive discussion of sources, and we illustrate interesting stylized patterns. We first document that cultural-ethnic factors play a primary role in the marriage market. We show, in particular, that gains from marriage are inversely related to spouses' cultural diversity. We then discuss further evidence suggesting that fertility and cultural-ethnic socialization of children might constitute important components of marriage gains. However, this analysis is necessarily silent with respect to the identification of cultural group preferences. It represents a motivation for the structural analysis of marriage, fertility, and socialization in the rest of the paper.

2.1 Homogamy and gains from marriage

We take advantage of administrative individual-level data from the Italian Statistical Institute (ISTAT, ADELE Laboratory), covering the universe of marriages formed in Italy from 1995 to 2012. The final sample counts more than 4 million marriages. In terms of cultural-ethnic traits, we distinguish between Italians, the native majoritarian group, and immigrant minorities, aggregated by country of origin.[7] We obtain $K = 7$ cultural-ethnic

[7]Specifically, first-generation immigrants are classified based on their country of origin, while second-generation immigrants are classified based on the country of origin of parents. Immigrants' cultural-ethnic

groups: n = Italian; i = European (EU15 countries), Other European,[8] North African and Middle-Eastern, Sub-Saharan African, East Asian and Latin American. This classification reflects both the prevalence of each ethnic-group in Italy and the relative cultural distance of countries with respect to Italy.[9] Geographically, we divide Italy in 20 distinct marriage markets, at the regional level. From these data, we recover the distribution of marriages, by cultural-ethnic group of spouses and by region; see Table C.2.[10]

Figure 1 reports the homogamy rates for ethnic minorities across markets, i.e., the fraction of the members of a specific cultural-ethnic group which marries homogamously per region. It documents strong positive assortative mating along cultural-ethnic lines for all cultural minorities, as the observed homogamy rates are significantly higher than those implied by random matching (corresponding to the 45-degree line in the figure). Interestingly, the data reveal heterogeneous degrees of assortativeness across groups; e.g., particularly high for Sub-Saharan African and East Asian minorities. We also document systematic differences in selection into marriage. Among minorities, the probability of marrying is as high as 63% for Latin American individuals, 59% for Sub-Saharan African, and it reduces to 46% for North African-Middle Eastern individuals, the group that is less likely to marry.

The homogamy and heterogamy rates we report clearly suggest that cultural-ethnic factors constitute an important determinant of marriage choices. However, these rates depend on the distribution of men and women by group in the marriage market under consideration. To control for both the uneven distribution of cultural-traits in the population and for potential gender imbalances, we estimate the *gains from marriage* along cultural lines, i.e., a measure of the utility gains associated to a marriage, with respect to the outside option of remaining single, for each marriage type (Choo and Siow, 2006).[11] Gains are computed

group is identified by nationality in years 1995 to 1997 and not by country of birth, because of data limitation. Our results are robust when we estimate the model excluding these initial years.

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

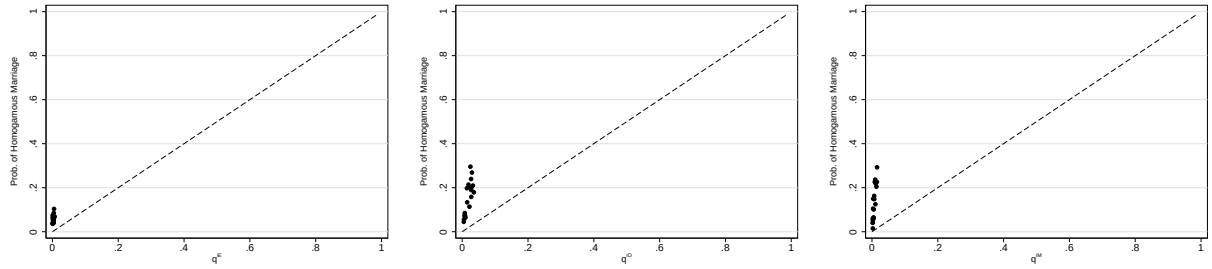
[9]Table C.1 reports the classification of foreign countries by cultural-ethnic group. Our classification aligns well with the main measures of cultural distance of foreign countries with respect to Italy, as documented in Figure C.1. In particular, our classification parallels the heterogeneity in genetic 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.

[10]Our data account only for legal marriages and only when celebrated in Italy. The increase in cohabiting couples is a very recent phenomenon. In 2011, the share of cohabiting native couples was still only about 9% in Italy - compared to the averages in the EU (14.5%) and the OECD (16.7%). Also, cohabitations are less relevant for our study since they are typically less stable and hence less motivated by fertility and children socialization; see (Lundberg and Pollak, 1993; Chiappori et al., 2017). Also, marriages celebrated out of the Italian marriage market are arguably less motivated by cultural integration issues.

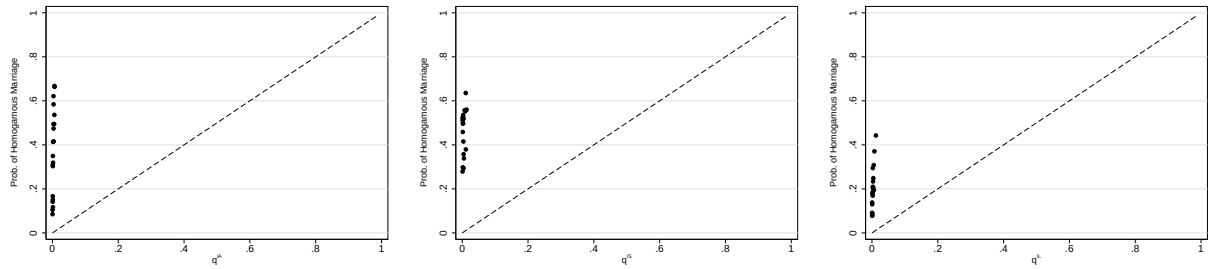
[11]Since Choo and Siow (2006), gains from marriage have been studied in different contexts of interest; see the references in Footnote 5.

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

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

by scaling the number of marriages for each cultural-ethnic type by the geometric average of the numbers of singles of those types. To this end, we complement marriage data with population Census data in 2001 and 2011 to obtain the distribution of single men and women by cultural-ethnic group and region.[12]

Table 1 reports average gains from marriage by marriage type.[13] Gains from marriage, like marriage rates, depend strongly on cultural-ethnic factors. Indeed, Table 1 shows strong positive assortative matching by cultural lines, with gains on the main diagonal being systematically larger than off-diagonal ones. For instance, the gains from assortative marriage across groups are especially high for Sub-Saharan African and East Asian minorities. The simple differences between homogamous and heterogamous families mask considerable ad-

[12]We select only adult single men and women (of more than 18 years of age); the distribution of single men and women is reported in Appendix Table C.3. To account for the possibility that single individuals might marry in the near future, we follow Chiappori et al. (2017); see Appendix A for details.

[13]The estimation procedure exposed in Table 1 implies that a high expected gain to marriage of a specific type (in terms of the cultural-ethnic groups of the spouses) is associated to high number of marriages of this same type, relative to the number of potential spouses of the same cultural-ethnic groups in the market. The estimated gains are negative, a common result in this literature, (see e.g., Choo and Siow, 2006).

ditional variation across cultural-ethnic groups and gender lines, which speak about the strength of the complementarity in cultural-ethnic traits.

We leverage this additional layer of variation, to study the relationship between gains from marriage and standard measures of cultural distance (Spolaore and Wacziarg, 2009, 2016) between the spouses. Indeed, in Panel A. of Table 4 we show that marriage gains are negatively correlated to the cultural distance of the ethnic traits of the spouses; this correlation is statistically significant - consistently - for different measures of cultural distance commonly used in the literature. We conclude from this analysis of homogamy and gains from marriage in our data that perceived cultural differences among groups are an important determinant of marriage allocations. We then turn to studying the mechanisms through which cultural differences affect marriage allocation.

2.2 Fertility and socialization rates

Cultural differences might affect marriage gains *per se*, conditioning the emotional and economic relationship between spouses. But, indirectly, cultural differences have also an effect on various components of marriage gains, notably on fertility and on children's socialization. Indeed we show that, more specifically, differential fertility and socialization rates by cultural-ethnic group could help account for the estimated distribution of marriage gains.

First of all, we provide evidence on fertility using matched registry data on marriages, and the universe of birth records for the period 1995-2012; see again Appendix A. Table 2 reports differential fertility rates between homogamous and heterogamous families, for all cultural-ethnic groups. Consistently with the hypothesised role of fertility as a component of marriage gains, homogamous marriages, which are associated to higher gains, display higher childbirth

Table 1: Gains from Marriage by Spouses Cultural-Ethnic Group

				Female Ethnic Group:			
Male	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah.Africa	East Asia	Latin America
Ethnic Group:							
Italian	-0.426	-4.184	-3.068	-6.488	-5.991	-5.857	-3.561
Europe-EU15	-4.502	-4.871	-6.812	-10.258	-9.061	-8.745	-7.465
Other Europe	-4.994	-7.439	-2.704	-10.009	-9.667	-9.292	-7.381
North Africa-Middle East	-5.804	-8.502	-6.198	-3.395	-9.814	-10.096	-7.083
Sub-Saharan Africa	-7.469	-9.057	-9.166	-10.812	-0.257	-11.375	-9.843
East Asia	-9.552	-10.552	-9.719	-11.432	-11.041	-1.035	-10.162
Latin America	-6.335	-9.132	-8.052	-11.239	-11.280	-10.742	-1.057

Notes: This table reports estimates for gains from marriage implied by the model. Consider a large marriage market, with a population of m men and f women, heterogeneous in terms of their cultural-ethnic identity. Let $\mu_{h,j}$ denote the fraction of marriages in the economy between a man of of cultural-ethnic identity h and a women j; let $\mu_{h\cdot}$ and $\mu_{\cdot j}$ denote the fraction of single men of trait h and single women of trait j, respectively. Gains from marriage G_{hj} are the utility gains associated to a hj marriage, with respect to the case in which both spouses remain single, for each hj marriage in each marriage market. Under standard assumptions which we make precise in the Appendix these gains are point identified from marriage data by type in each marriage market. We estimate gains from marriage, G_{hj}, by cultural-ethnic group of spouses h, j, as follows:

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

Table 2: Fertility Rates by Ethnic Group of Spouses

	Panel a. Extensive Margin Probability of Having a Child			Panel b. Intensive Margin Number of Children		
	Homogamous	Heterogamous	Heterogamous Italians excluded	Homogamous	Heterogamous	Heterogamous Italians excluded
Italian	0.735	0.418	-	1.561	1.390	-
Europe-EU15	0.420	0.578	0.359	1.465	1.490	1.385
Other Europe	0.494	0.401	0.381	1.293	1.365	1.342
North Africa-Middle East	0.541	0.303	0.350	1.462	1.297	1.313
Sub-Saharan Africa	0.534	0.352	0.247	1.592	1.369	1.346
East Asia	0.682	0.308	0.282	1.516	1.302	1.298
Latin America	0.322	0.344	0.359	1.241	1.308	1.295

Notes: This table reports fertility rates by ethnic group of spouses, both in terms of the probability of having children (panel a) and in terms of the average number of children (panel b). Estimates are reported separately for homogamous, heterogamous, and heterogamous families excluding marriages with natives.

investments, both at the extensive and intensive margin; see also Figure C.2. For example, the probability of having a child in a heterogamous marriage with at least one Italian spouse is of 41.8%, compared to the 73.5% of homogamous native marriages. As marriage gains, fertility also displays significant variation across cultural-ethnic groups: at the extensive margin, for instance, the probability of having a child in a heterogamous marriage is 58% for Europeans, while only 31% for East Asians. Importantly, this heterogeneity is at least partly explained by cultural differences. Panel B of Table 4 shows that fertility is indeed negatively correlated to different measures cultural of distance.[14]

With regards to socialization, we focus on language socialization. We recover socialization probabilities by cultural-ethnic group of spouses and by region from the *Condition and Social Integration of Foreign Nationals* Survey (2011-2012). The survey is targeted to foreign residents in Italy and it is intended to provide a comprehensive representation of their socio-cultural as well as economic integration. 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/she is socialized to his/her mother language, defined as the idiom acquired during the preschool period of childhood. We interpret speaking Italian at home as a form of parental investment in integration, and a signal of lower preferences for cultural identity. To support our interpretation, in Section 5.3.2, we provide consistent evidence that speaking Italian at home is, in fact, associated with higher achievement and educational choices of immigrant students, weaker ethnic identity, and stronger attitudes towards social integration.

[14]Interestingly, we document similar patterns in terms of marital instability (separations) of heterogamous marriages as compared to culturally homogeneous unions; see Appendix Table C.5.

Table 3: Fraction of "Italian spoken at home" by Ethnic Group of Spouses

	Italian Socialization Probabilities		
	Homogamous	Heterogamous	Heterogamous Italians excluded
Italian	1	0.915	-
Europe-EU15	0.442	0.867	0.641
Other Europe	0.395	0.925	0.803
North Africa-Middle East	0.267	0.884	0.706
Sub-Saharan Africa	0.377	0.891	0.727
East Asia	0.196	0.816	0.692
Latin America	0.469	0.904	0.926

Notes: This table shows socialization outcomes by ethnic group of spouses. The outcome variable is an indicator for whether the child speaks Italian within the family. Estimates are reported separately for homogamous, heterogamous, and heterogamous families excluding marriages with natives.

Table 3 reports the fraction of "Italian spoken at home" by cultural-ethnic group for homogamous and heterogamous families. Consistently with socialization being also a component of marriage gains, besides fertility, the table documents high socialization rates (low rates of "Italian spoken at home"), especially for homogamous marriages of minorities: the probability that an immigrant parent speaks Italian with his/her child in homogamous marriages is generally half as high as in heterogamous marriages with another immigrant, and higher in heterogamous marriages in general; see also Figure C.3.[15] The variation across cultural-groups is also high: e.g., the probability that an East Asian parent speaks Italian with his child is equal to 20% in a homogamous marriage (82% in a heterogamous marriage), while this probability is 44.4% for a European (86.7% in a heterogamous marriage). Interestingly, the relationship between socialization rates and cultural distance in heterogamous marriages, reported in Panel C of Table 4, has an opposite sign, though not significant: the larger the cultural distance the higher socialization.

The evidence we described regarding fertility and socialization as 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, emphasizing the central role of cultural-ethnic socialization of children in marriage.[16] Immigrant parents wishing to preserve their cultural traits might have a preference for marrying homogamously or at least

[15]This is consistent with other evidence, using different measures of socialization (among others Dohmen et al., 2012; Fouka, 2020). The fraction of "speaking Italian at home" in divorced homogamous households is higher than in married ones; see Table C.4. This suggests that the impact of divorce on socialization might also play a role on marriage choices; we formalize this link in Section 3 and estimate it in Section 4.

[16]See e.g., the fundamental work of Boas (1928); Lévi-Strauss (1949) in anthropology; see also Riesman (1992); Smith (1996); Mayer (2013), and, in economics, Bisin and Verdier (2000); Bisin et al. (2004).

with spouses with close cultural-ethnic traits, to limit their children's exposure to culturally distant traits. Similarly, native parents might also have a preference to limit their children's exposure to other cultural-ethnic traits, in heterogamous marriages. In this case, adopting a quantity-and-quality of children's metaphor for illustration purposes (Becker, 1973, 1974), marriages between culturally close spouses are associated with high gains in that they generate high quality children, inducing in turn high fertility (quantity).

Following this interpretation of our analysis, the negative association of cultural distance with marriage gains and children quantity is fundamental in ordering our understanding of marriage and fertility choices along cultural lines. On the other hand, this analysis falls short of allowing us to identify several fundamental aspects of cultural preferences in the marriage market, that is, e.g., to address how much the observed marriage allocation is determined by the immigrants' preferences to preserve their cultural identity or by different forms of the natives' (lack of) acceptance of immigrants' cultural-ethnic traits. Indeed, any measure of cultural distance is necessarily symmetric,[17] while arguably the distribution of preferences along the cultural-ethnic dimension is not, between immigrants and natives but also differentially with respect to the cultural-ethnic traits of immigrants. Furthermore, the cultural distance associated to any homogamous marriage is null by construction, independently of the cultural-ethnic group, making it impossible to account for the significant observed variation of the distribution of homogamy by group.[18]

Motivated by this empirical investigation, we turn to study marriage gains, socialization, and fertility explicitly as the outcome of an equilibrium process of household formation and intra-household cultural decisions. In the rest of the paper, we develop and structurally estimate an equilibrium model of marriage, fertility, and socialization, and identify the distribution of cultural preferences along the cultural-ethnic dimension. The model allows us to study the integration patterns of the different cultural-ethnic minorities in Italian society.

3 A model of marriage, fertility, and socialization

Consider a frictionless marriage market. 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.[19] Transfers are endogenously determined as

[17] For recent advancements, see Vieira et al. (2022).

[18] Indeed, the observed non-significant but positive relationship between socialization rates and cultural distance in Panel C of Table 4, suggests a more complex view of socialization, whose intensity might depend on the relative strength of the immigrant parents' preferences to preserve their culture.

[19] 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.

Table 4: Correlation of Marriage Outcomes and Cultural Distance Measures

	Panel A. Dep. var: Gains from Marriage			
Genetic Distance	-5.297***			
	(0.324)			
Linguistic Distance		-5.970***		
		(0.139)		
Religious Distance			-6.297***	
			(0.216)	
Cultural distance index, WVS				-7.689***
				(0.455)
Observations	628	628	628	628
R-squared	0.348	0.582	0.626	0.458
	Panel B. Dep. var: Fertility rates			
Genetic Distance	-0.428***			
	(0.047)			
Linguistic Distance		-0.558***		
		(0.042)		
Religious Distance			-0.605***	
			(0.053)	
Cultural distance index, WVS				-0.808***
				(0.094)
Observations	628	628	628	628
R-squared	0.212	0.475	0.541	0.473
	Panel C. Dep. var: Italian socialization Rates			
Genetic Distance	-0.012			
	(0.043)			
Linguistic Distance		-0.163		
		(0.293)		
Religious Distance			-0.040	
			(0.068)	
Cultural distance index, WVS				-0.039
				(0.060)
Observations	306	306	306	306
R-squared	0.122	0.123	0.124	0.124

Notes: This table shows OLS estimates of the relationship between marriage and intra-household outcomes and cultural distance. The outcomes are gains to marriage, computed as in equation (1), in Panel A; fertility and socialization rates in Panel B and C, respectively. 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 details about the variable construction and definitions. Regressions are weighted by province population. Robust standard errors clustered at regional level in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

equilibrium outcomes, depending on the quality of the specific match but also on the set of available opportunities in the marriage market.

Family types are heterogeneous. The notation $\{h, j\}$ denote a household type where the male has cultural-ethnic identity h and the female j; let $\{h, .\}$ denote the household type

composed of a single male with trait h and $\{., j\}$ the type composed of a single female of type j. Let T denote the set of possible types of household, including those composed of single individuals, along the cultural-ethnic identity. Abusing notation, we use $t \in T$ to index all types of household, and $hj \in T$ to index married household. For simplicity, we present the theoretical model for dichotomous cultural traits; say n for natives and i for immigrants, while the empirical analysis allows for multiple cultural-ethnic traits of immigrants.

Total marital utility is the sum of two components: i) a systematic component, and ii) an idiosyncratic component, capturing residual idiosyncratic returns from marriage, observed by the individuals prior to marriage. Let ϵ_h and η_h denote the individual idiosyncratic preference shocks for men and women, respectively, with identity h; a vector, each element of which represents the idiosyncratic component of utility associated to a possible type of spouse the individual might be matched with (including none, if he/she stayed single). The total expected utility of a household of type hj between man m with identity h and female f of identity j is assumed additive and separable in the shocks, $U_{hj} + \epsilon_{hj} + \eta_{jh}$ (resp. $U_{h.} + \epsilon_{h.}$). In the marriage market, individuals observe their idiosyncratic shocks and match along cultural-ethnic identity traits, anticipating their marital utility $U_t + \epsilon_t + \eta_t$ for all different potential matches.

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

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

where $\varepsilon(\mu)$ represents the generalized entropy of the matching, which means to capture the dispersion of individual preferences with respect to the aggregate preferences, conditional on spouses' attributes (Galichon and Salanié, 2017, 2021). Following Choo and Siow (2006), we assume that ϵ_t, η_t are independent and identically distributed random variables with a type I extreme-value distribution (Gumbel). The matching model in (2) translates into a

multinomial logit model (McFadden, 1974).[20] As a consequence, equilibrium gains G_{hj},[21] whose estimates we reported in 1, satisfy

$$G_{hj} = \log \frac{(\mu_{hj})^2}{\mu_{h.}\,\mu_{.j}}. \tag{4}$$

We further postulate that the systematic utility of marriage U_{hj} has two components:

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

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

$$U_{hj}^{ec} = \alpha\,(U_{h.} + U_{.j})\,;$$

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

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

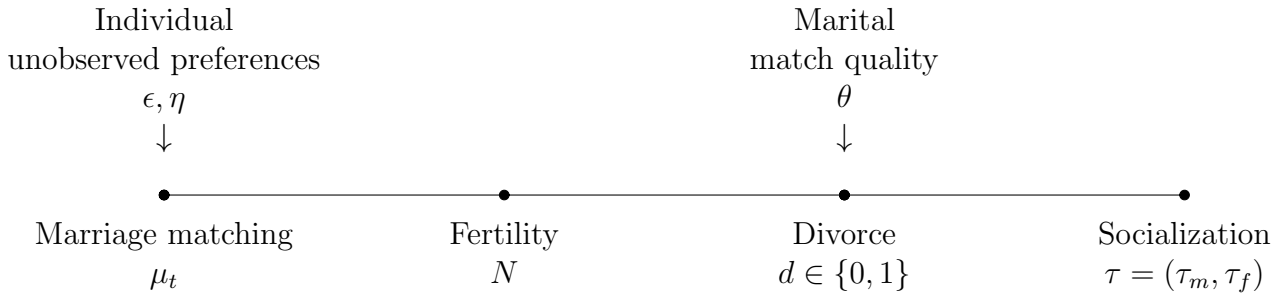
[20] Under this separability assumption, Galichon and Salanié (2021) and Chiappori et al. (2017) specifically show that the two-sided matching problem reduces to a series of one-sided discrete choice problems. From Galichon and Salanié (2021), the generalized entropy in (2) is:

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

where $T\backslash\{.\} = \{n, i_E, i_O, i_M, i_A, i_S, i_L\}$, that is, the set T once singlehood $\{.\}$ is excluded.

[21] Formally, $G_{hj} = U_{hj} + \epsilon_{hj} + \eta_{hj} - U_{h.} - \epsilon_{h.} - U_{.j} - \epsilon_{.j}$.

Figure 2: Timing of the Model



cooperative decision by both parents in the household or as a non-cooperative decision of the mother in case the household is separated.[22] We consider a series of simplifications. First of all, the marital utility is proportional to fertility N and investment in fertility entails a cost $\kappa(N)$, increasing and convex in N. Secondly, the utility per child is composed of i) a utility from socialization $u_{hj}(\theta)$;[23] and ii) a direct utility δ from having a child in the marriage as opposed to outside the marriage. Furthermore, the marital quality shock θ enters marital utility only if the household stays married (does not divorce); that is, if it chooses $d = 0$:

$$U_{hj}^{par} = N\left[E\left(u_{hj}(\theta)\right) + \delta(1 - d(\theta))\right] + E\left(\theta(1 - d(\theta))\right) - \kappa(N). \tag{5}$$

From the above equation notice that the systematic value of a match arises endogenously in the model, and it is a function of the (match specific) expected utility that parents derive from socialization $E\left(u_{hj}(\theta)\right)$.

We proceed backwards, to introduce in more detail the various components of U_{hj}^{par} from socialization to fertility.

Socialization. We start from the socialization problem, given (N, θ, d). In fact, under the preference structure we imposed, the socialization choice is independent of fertility N and it depends on θ only through d. Let V_j^h denote the utility a parent with trait j obtains if the child is socialized to trait h, for all h, j. Each parent's preference over the cultural-ethnic identity of his/her children is biased towards his/her own trait, as a manifestation of paternalistic altruism:

[22] 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.

[23] More precisely, $u_{hj}(\theta) = u_{hj}(\tau, d)$, where $\tau = \tau(d)$ and $d = d(\theta)$.

$$V_h^h > V_h^j, \text{ for all } h \neq j \in \{n, i\}.^{24}$$

With regards to the socialization technology, following Cavalli-Sforza and Feldman (1981), we interpret the process of transmission of cultural traits as the interaction of two forces: the *vertical* socialization of parents within the family, and the *horizontal* socialization of the society at large. As for the vertical socialization at the level of the family, we introduce several simplification assumptions. First of all, within a family all children identify to the same trait.[25] Secondly, homogamous native households socialize their children with probability 1; that is, children of native parents speak the native language. Thirdly, in a household of type hj the socialization effort of the father, τ_m, has the objective and the effect of increasing the probability that the children identify with his trait, h; 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, j. Parents in heterogamous households, such that $h \neq j$, face conflicting incentives in the socialization of children, while parents in homogamous households, with $h = j$, benefit from coordinated incentives. Thus, the value of the marriage derives from the coordination of investments in children. Finally, the socialization technology responds to the social environment. Let q^h define the fraction of individuals with trait h 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^h.

Let $P_{hj}^h(\tau, d)$ denote the probability that a child in a family of type hj is socialized with the father's trait $h = n, i$, 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, $P_{nn}^n(\tau, 0) = P_{nn}^n(\tau, 1) = 1$. Assuming that the mother is given custody of children in divorce, socialization technologies, extending Bisin and Verdier (2000), are as follows:

$$\begin{array}{ll} P_{ii}^i(\tau,0) = \tau_m + \tau_f + (1-\tau_m-\tau_f)q^i, & P_{ii}^n(\tau,0) = (1-\tau_m-\tau_f)(1-q^i) \\ P_{in}^i(\tau,0) = \tau_m + (1-\tau_m-\tau_f)q^i, & P_{in}^n(\tau,0) = \tau_f + (1-\tau_m-\tau_f)(1-q^i) \\ P_{ni}^i(\tau,0) = \tau_f + (1-\tau_m-\tau_f)q^i, & P_{ni}^n(\tau,0) = \tau_m + (1-\tau_m-\tau_f)(1-q^i). \end{array} \tag{6}$$

Socialization probabilities under divorce $P_{hj}^h(\tau, 1)$ are equivalent to those reported in (6),

[24] By assuming that V_h^h and V_h^j are constant parameters, independent of economic conditions, we do not distinguish cultural and economic incentives for socialization; see Bisin and Verdier (2001, 2011).

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

after imposing $\tau_m = 0$. The total marital utility from the socialization process, net of increasing and convex socialization costs $c(\tau)$, is:

$$\begin{aligned} u_{hj}(\tau, d) &= P_{hj}^h(\tau, d)\left(V_h^h + V_j^h\right) + \left(1 - P_{hj}^h(\tau, d)\right)\left(V_h^j + V_j^j\right) - c(\tau) \\ &= V_h^j + V_j^j + P_{hj}^h(\tau, d)\left(\Delta V_h^j - \Delta V_j^h\right)\mathbf{1}\left(\Delta V_h^j > \Delta V_j^h\right) \\ &\quad + P_{hj}^j(\tau, d)\left(\Delta V_j^h - \Delta V_h^j\right)\mathbf{1}\left(\Delta V_j^h \geq \Delta V_h^j\right) - c(\tau). \end{aligned}$$

where $\Delta V_h^j = V_h^h - V_h^j$ and $\Delta V_j^h = V_j^j - V_j^h$ are referred to as the *cultural intolerance* of cultural-ethnic group h and j, respectively.

Socialization effort τ is then the solution to

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

Let the solution be denoted $\tau(d)$.[26] Notice that it depends only on ΔV_h^j, ΔV_j^h rather than on the utility levels V_j^h (which do not affect the maximization problem in (7)). Moreover, at the solution, the parents' choice of socialization effort is also a function of q^i, i.e., of the proportion of immigrants of group i in the reference population.

Divorce. After observing the realization of the marriage quality shock θ, the spouses optimally choose whether to dissolve the marriage (divorce) or not, rationally anticipating their total utility from the socialization process. Given N, a type hj household divorces, choosing $d(\theta) = 1$, if

$$N\left(u_{hj}(\tau(1), 1)\right) > N\left(\delta + u_{hj}(\tau(0), 0)\right) + \theta.^{27}$$

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

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

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

[26] Whenever possible without confusion, we avoid to use the hj subscript in the notation.
[27] Notice that θ is a match (not individual) specific random variable.

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

3.1 Results

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

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

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

For all household types, married or divorced, the probability of successful socialization to the trait desired by the parents (or parent) doing the investment is greater than the rate associated to horizontal socialization.[28] We turn now to study comparative statics,

In homogamous minority households, whether parents divorce or stay married, both parents' socialization efforts are monotonically increasing in cultural intolerance and decreasing in the size of their cultural group, q^i. In heterogamous households, the socializing parent effort is monotonically increasing in his/her own cultural intolerance; if parents stay married, the socializing parent's effort is also decreasing in his/her spouse's cultural intolerance. It is also the case that the minority i socializes more than the majority n, ceteris paribus.[29]

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

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

Divorce. Consider an household with positive fertility, $N > 0$. As the systematic gains from marriage derive from socialization, and divorce leads to a generally less efficient socialization technology,

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

On the other hand, in heterogamous households, mothers have an advantage in socialization after divorce,[30]

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

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

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

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

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

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

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

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

4 Structural estimation: Methodology

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

4.1 From the model to the data

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

Concerning socialization probabilities, we assume that in households ii, in and ni children can only be socialized either to trait i or n; while in a heterogamous household with both immigrants parents, the children can be socialized either to one of the parents' traits or to n.[31] The remaining socialization probabilities are constrained to be zero. Also, we allow socialization and fertility costs to capture systematic differences between homogamous and heterogamous couples, indexed by $s \in \{het, hom\}$, respectively:

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

where $\xi_s \geq 1$, captures the dependence of fertility costs on childbearing decisions.[32]

For flexibility in the estimation, we allow the distribution of θ_{hj} to have a mean that depends on the household type hj. More specifically, we assume that θ_{hj} follows a generalized logistic distribution with location a_{hj} and scale parameter b. We normalize $b = 1$ and we set a_{hj} to match the dissolution probability of couples *without children* in the data for all hj;

[31] Specifically, in a generic hj heterogamous marriage with $h \neq j$ and $h, j \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$, with $d = 0$, the socialization rates to Italian, P_{hj}^n, to the father's language, P_{hj}^h, and to the mother's language, P_{hj}^j, are respectively: $P_{hj}^n(\tau, 0) = (1 - \tau_m - \tau_f) q^n$; $P_{hj}^h(\tau, 0) = \tau_m + (1 - \tau_m - \tau_f) q^h$; $P_{hj}^j(\tau, 0) = \tau_f + (1 - \tau_m - \tau_f) q^j$.

[32] 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.

i.e., $a_{hj} : F(0; a_{hj}, b) = \hat{\pi}_{hj}(0)$.[33] This assumption allows us to capture systematic differences in separation rates across household ethnic groups without children; that is, independently from children socialization mechanism. Aggregate evidence is reported in Table C.6.

We allow the residual value of marriage, i.e., the value of staying single, to vary with the ethnic group and separately for homogamous and heterogamous marriages, ω_{h_s}, for $h \in \{n, i\}$ and $s \in \{hom, het\}$. This is to capture, indirectly, differential sorting in both observables and unobservables across cultural-ethnic groups and between families.

Finally, we model the role played by the immigrants' cultural-ethnic network within the transmission process, relaxing the initial assumption of unbiased horizontal socialization frequencies.[34] A strong network fosters the ability of immigrants' communities to pass on their cultural identity to new generations with no direct investments from parents (e.g., public housing, freedom of religion, schooling). In our framework, this is captured indirectly by introducing a positive segregation bias, ρ, allowing each minority i to face a segregated socialization pool composed of a fraction Q^i of individuals of the same group i; where

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

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

4.2 Parameters and Estimation

The main parameters of interest are the cultural intolerance parameters, $\Delta V_h^j = V_h^h - V_h^j$, for all cultural-ethnic groups h, j. In the estimation, we impose V_h^h to be constant across groups, for all h ($V_h^h = V$); that is, we assume that the value for a parent in sharing the same cultural trait of his/her own child is constant across groups, while relative differences in transmission are allowed to vary across groups. For identification purposes, we normalize $V = 100$, so that cultural intolerances are measured in units corresponding to percentages

[33] 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.

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

of V. We are left with $K(K-1) = 42$ cultural intolerance parameters to estimate. The other parameters to be estimated are: socialization and fertility cost function parameters, $\sigma_{\tau_s}, \lambda_{\tau_s}$ and $\sigma_{N_s}, \lambda_{N_s}$; dependence of fertility costs on childbearing decisions, ξ_s; direct value of fertility (independently from cultural socialization), δ; segregation bias ρ; outside option of being single ω_{h_s}, for all h and s; and relative effects of marriage on spouses' economic opportunities, α.

Let $\boldsymbol{\beta}$ denote the vector of parameters. Given an exogenous population distribution q^h, for all groups h, the structural model provides us with the theoretical moments in reduced form, $\widetilde{\mathbf{\Pi}}(\boldsymbol{\beta})$. Specifically, in our estimation, the theoretical moments we exploit are maps from $\boldsymbol{\beta}$ into $\tilde{U}_{hj}$, N_{hj}, π_{hj}, for all hj, and $P^k_{hj}(d)$ for all hj and k, and marital status d.[35]

The empirical moments are $\widehat{\mathbf{\Pi}} = \{\hat{U}_{hj}, \hat{N}_{hj}, \hat{\pi}_{hj}, \hat{P}^k_{hj}\}$, 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 (4), where $\hat{\mu}_{hj}$ is obtained from the distribution of marriages over the period 1995-2012, while $\hat{\mu}_{h.}$ and $\hat{\mu}_{.j}$ are taken from the population vectors by ethnic group, gender and marital status of individual Census data in 2001 and 2011. We compute fertility rates $\hat{N}_{hj}$ as the average number of children in households of type hj, including zeros.[36] We evaluate separation rates $\hat{\pi}_{hj}$ as the fraction of marriages of type hj ending in separation during the period of analysis, conditional on having children. Finally, we construct socialization frequencies, $\hat{P}^k_{hj}$, as the fraction of households of type hj in which children speak a given language k at home.[37] Given normalization restrictions, we end up with a total of 69 parameters to match 2,416 moments.[38]

We estimate model parameters via a method of moments estimator, by matching the vector of theoretical moments implied by the model, $\widetilde{\mathbf{\Pi}}(\boldsymbol{\beta})$, for a specified choice of parameters $\boldsymbol{\beta}$, with their empirical counterparts observed in the data, $\widehat{\mathbf{\Pi}}$. Formally, given a weighting

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

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

[37] 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.

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

matrix $\mathbf{\Omega}$,[39]

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

4.3 Identification

Our estimation procedure exploits two sources of cross-sectional variation: variation across cultural-ethnic groups and family types as well as variation in the ethnic composition of the population across regions. Identification, thus, requires us to assume that all parameters are constant across the 20 geographical regions $r \in R$ and that each region corresponds to a separate local marriage market.[40] Furthermore, identification hinges also on i) the random variable θ_{hj} having the same distribution across households ethnic groups hj; ii) the segregation bias, ρ, the economic complementarity in marriage, α, as well as the socialization and fertility cost parameters $(\sigma_{\tau_s}, \lambda_{\tau_s}, \sigma_{N_s}, \lambda_{N_s}, \xi_s, \delta)$ being independent across ethnic groups h. Independence of costs, in particular, implies that any difference in costs across groups would be attributed, in our estimates, to cultural intolerances. On the other hand, we can allow outside options of being single, ω_{h_s}, to be specific to the ethnic group h and differ by family type s. These parameters are pinned down by the average probability of marrying for each ethnic group and family type across regions.

Under these assumptions, no restrictions need be imposed on cultural intolerance parameters: geographical variation in population vectors allows us to separately identify the cultural intolerance of parents of type h with respect to children of type j, ΔV_h^j, from the intolerance of parents of type j with respect to children of type h, ΔV_j^h. Furthermore, and most importantly, we can identify the cultural intolerance parameters of minorities with respect to children integrated as natives, ΔV_i^n, separately from cultural intolerances of natives versus all minorities, ΔV_n^i, for all i. This is, in fact, tantamount to identifying demand and supply components of cultural integration as an equilibrium phenomenon. Such identification is possible in our setup because theoretical moments for homogamous immigrant families are a function of demand parameters only, and thus identify demand. Conditional on demand, theoretical moments of heterogamous marriages with natives pin down supply parameters. We should note, however, that while we allow cultural intolerance parameters

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

[40] We calculate that in more that 92% of our marriages, spouses share the same region of residence.

to be cultural-ethnic group specific, we need to assume they are constant across households within cultural-ethnic group.[41] Within group variation in cultural intolerance would imply that individuals with lower intolerance would more frequently marry with natives: our estimates would underestimate the fraction of intermarriages, and we expect they would underestimate socialization and fertility costs as well as outside options.

5 Structural estimation: Results

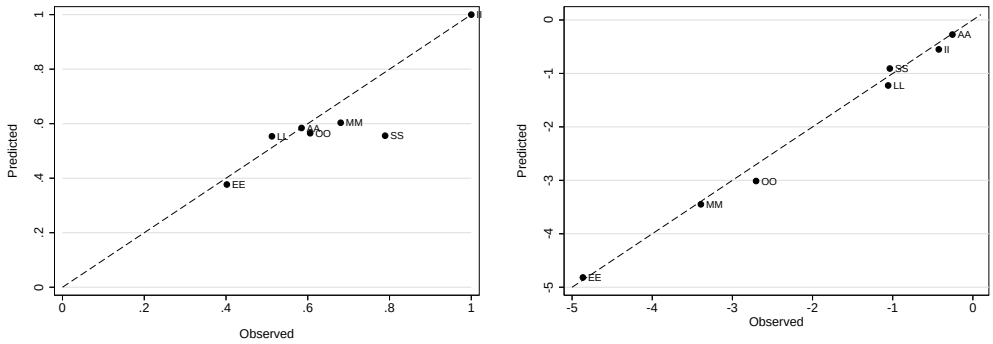
In this section, we start by describing the fit of the model. We then introduce the parameter estimates. Finally, we present various validation exercises and we discuss the main assumptions of our empirical model, reporting on several robustness checks. As for the implication of our estimates with regards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

5.1 Model Fit

The model fits the data well. The raw correlation between predicted and implied gains from marriage from the model is equal to 0.84 and the marriage patterns observed in the data are matched very well in our empirical exercise; see Figure C.5. Table 5 compares the average observed and predicted moments, for homogamous and heterogamous families,

Figure 3: Fit of the Model in Homogamous Families - Socialization Rates and Gains to Marriage

(a) Socialization rate to native culture (b) Gains to Marriage



Notes: This figure shows the average fit of the model by household for homogamous families, considering socialization rates to native culture (panel a) and gains from marriage (panel b). I: Italians; E: Europe-EU15; O: Other Europe; M: North Africa-Middle East; A: Sub-Saharan Africa; S: East Asia; L: Latin America.

[41] In addition, we need to assume that preference parameters are constants along gender lines. In principle, we would have introduced gender differences in the outside options of being single, ω_{h_s}.

respectively. Overall, we match well the socialization frequencies with a correlation between predicted and observed foreign language socialization rates of 0.83, for both homogamous and heterogamous families. Figure 3 displays the average fit for the rate of socialization to the native culture and the gains to marriage in homogamous families. Similarly, the model fits well the fertility rates for heterogamous families and for some homogamous families, but less so for some others. The model is also able to capture the general pattern of separation choices across groups, even though separation rates appear to be slightly underestimated. Finally, the model is able to capture the geographical variability across markets. Focusing on homogamous immigrant families, Figure C.6 displays the relationship between predicted and implied gains from marriage over the corresponding population share by region of residence.

Table 5: Fit of the Model

	a. Homogamous Families									
	Italian Soc		Father Soc		Separation		Fertility		Marital Gain	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Italian	1.00	1.00	1.00	1.00	0.05	0.06	1.12	1.08	-0.43	-0.55
EU15	0.60	0.62	0.40	0.38	0.03	0.01	0.69	1.00	-4.87	-4.82
Other Europe	0.39	0.44	0.61	0.56	0.02	0.01	0.64	0.84	-2.70	-3.01
North Africa-Middle East	0.32	0.40	0.68	0.60	0.02	0.02	0.79	0.69	-3.39	-3.45
Sub-Saharan Africa	0.42	0.42	0.58	0.58	0.02	0.01	0.85	0.74	-0.26	-0.27
East Asia	0.21	0.44	0.79	0.56	0.01	0.01	1.06	0.83	-1.04	-0.91
Latin America	0.49	0.45	0.51	0.55	0.03	0.02	0.40	0.83	-1.06	-1.23
	b. Heterogamous Families									
	Italian Soc		Father Soc		Separation		Fertility		Marital Gain	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Italian	0.95	0.93	0.69	0.67	0.05	0.03	0.58	0.52	-4.16	-4.07
EU15	0.93	0.97	0.50	0.53	0.04	0.02	0.85	0.62	-4.53	-4.45
Other Europe	0.96	0.95	0.75	0.74	0.04	0.02	0.55	0.50	-3.63	-3.58
North Africa-Middle East	0.95	0.93	0.31	0.32	0.07	0.07	0.39	0.38	-6.17	-6.16
Sub-Saharan Africa	0.91	0.97	0.63	0.62	0.06	0.04	0.48	0.43	-6.70	-6.56
East Asia	0.87	0.95	0.82	0.80	0.04	0.01	0.40	0.50	-6.62	-6.66
Latin America	0.93	0.71	0.81	0.61	0.05	0.05	0.45	0.52	-4.24	-4.23

Notes: This table shows the fit of the model by cultural-ethnic group of spouses, separately for homogamous (panel a) and heterogamous (panel b) families. Estimates are weighted by the number of marriages by match and region. In addition, see Figure C.7 for a graphical representation.

5.2 Parameter estimates

Table 6 presents the estimation results.[42] Cultural intolerance parameters are estimated strictly greater than zero; that is, parents of each cultural-ethnic group have preferences for socializing children to their own cultural-ethnic group. Second, intolerance parameters are highly heterogeneous across cultural-ethnic groups, i.e., some groups are much more resilient in their cultural identity compared to others. Notably, e.g., immigrants from North Africa-Middle East have very strong preferences for maintaining their cultural identity: A child integrated to the native culture, for a North African parent, is valued 65% less than one socialized to culture of the parent. For a parent of a EU15 country this loss is only about 10%. The cultural intolerances of Italian natives are also heterogeneous towards different minorities. To an Italian parent, a child socialized to the Sub-Saharan African cultural traits implies a 78% loss, a much larger loss than if socialized to Latin American traits, about 20%.

Third, the matrix of intolerance parameters is largely asymmetric, i.e., the intolerance of group h versus group j is often not quantitatively close to the intolerance of group j versus group h; see Figure 4. Notably, e.g., natives appear particularly accepting of Latin American immigrants, as we already noted; while the intolerance of Latin Americans towards

Table 6: Structural Model Parameters

Cultural Intolerance Parameters							
h:	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah Africa	East Asia	Latin America
ΔV_n^h, Italian		33.38	58.60	67.88	78.23	61.53	19.27
$\Delta V_{i_E}^h$, Europe-EU15	10.21		52.60	4.77	6.69	18.54	0.33
$\Delta V_{i_O}^h$, Other Europe	39.97	0.05		69.32	56.63	30.42	23.02
$\Delta V_{i_M}^h$, North Africa-Middle East	65.35	7.00	58.98		97.85	52.12	52.25
$\Delta V_{i_A}^h$, Sub-Saharan Africa	55.00	28.37	58.42	96.37		81.22	42.50
$\Delta V_{i_S}^h$, East Asia	40.02	0.30	87.08	54.70	47.29		93.13
$\Delta V_{i_L}^h$, Latin America	38.95	10.69	20.77	58.99	46.96	29.13	
Outside Option of Singlehood Parameters							
Outside option for homogamous, ω_h	83.39	72.34	44.45	28.58	30.93	41.61	41.74
Outside option for heterogamous, ω_h	39.67	54.19	19.67	3.65	12.44	25.65	25.08
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					

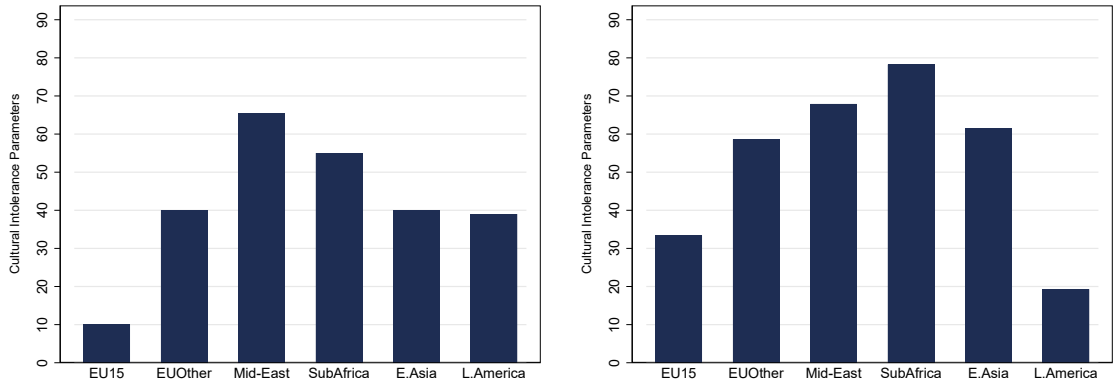
Notes: This table shows structural parameter estimates.

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

Figure 4: Cultural Intolerance Parameters

(a) Immigrants towards Natives (Demand parameters)

(b) Natives towards Immigrants (Supply parameters)



Notes: This figure reports parameter estimates for the cultural intolerance of immigrants versus natives ΔV_i^n (panel a) and natives versus immigrants ΔV_n^i (panel b) for all cultural-ethnic minorities i.

natives is twice as large. The low estimated intolerance of natives towards Latin American immigrants is required to fit the high number of intermarriages of Italian men with Latin American women and the high fertility rate in these intermarriages (similar in magnitude to the fertility rate of homogamous Latin American marriages).

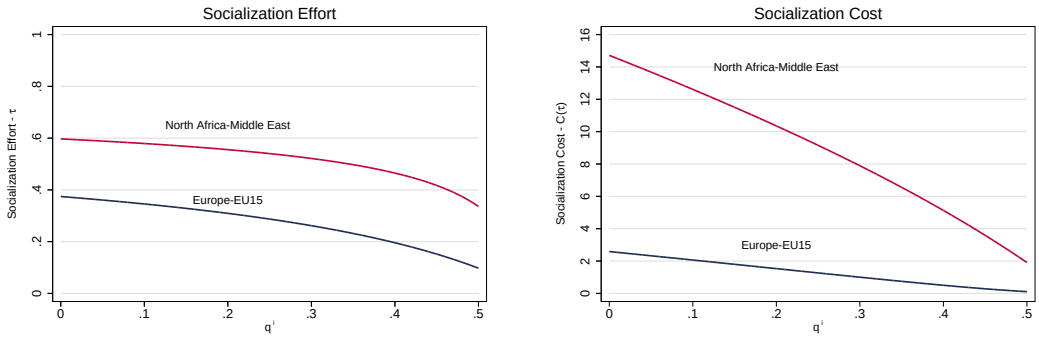
These differences is cultural intolerance preferences across groups also translate into significant differences in socialization investments. For instance, the direct socialization effort of homogamous families, τ, calculated in the extreme case in which the family belongs to a full minority i ($q^i = 0$), is estimated between 0.38 (Europe-EU15) and 0.60 (North Africa-Middle East), i.e., North Africa-Middle East parents have 50% higher probability of socializing children directly to their own culture, compared to Europe-EU15 parents in homogamous families. Figure 5 displays the direct socialization investment implied by our estimates, for Europe-EU15 and North Africa-Middle East (the two groups with extreme patterns), as a function of the fraction of their group in the population. For homogamous families (panel a), τ^i declines with q^i; that is, families substitute between vertical and horizontal socialization which is a consequence of the children's social interactions in the population at large. Figure C.8 reports the implied socialization probabilities. Heterogamous families with natives (panel b), instead, do not directly socialize children to the immigrant parent's culture, and the socialization effort of the native spouse τ^n increases with q^i.

Socialization costs across families are significantly different in our estimates. The cost σ_τ of heterogamous families is twice as high as the one of homogamous families.[43] This difference

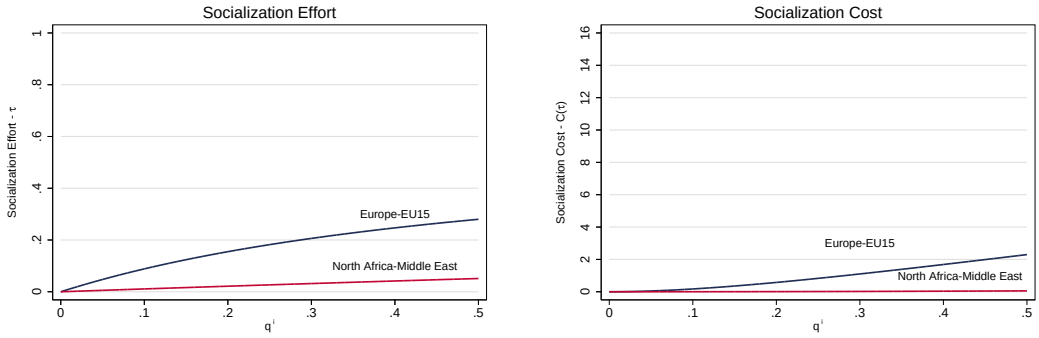
[43]Costs functions are assumed to be independent of the cultural-ethnic group of spouses. The estimates

Figure 5: Estimates of Socialization Effort and Socialization Cost by Minorities

(a) Homogamous



(b) Heterogamous



Notes: This figure reports estimates of socialization effort, τ, and the corresponding socialization costs, $c(\tau)$, over the potential population share, q^i, for European-EU15 and North African-Middle East minorities. Panel a. reports estimates for homogamous families, panel b. reports estimates for heterogamous marriages with natives.

in socialization costs reinforces the gap in socialization investments between family types. The corresponding socialization cost functions, $c(\tau)$, are described in Figure 5. We also estimate a positive difference in fertility cost σ_N, one third greater for heterogamous than for homogamous families.[44] Overall, our estimates imply that fertility investments are much more costly whenever spouses belongs to different cultural groups.

The parameters capturing the outside option of remaining single are estimated to be highly heterogeneous both across families (homogamous vs heterogamous) and across cultural-ethnic groups, with homogamous natives showing the highest outside option parameters and the heterogamous North Africa-Middle East group showing the smallest ones. Finally, we estimate a segregation bias, ρ, of about 2; that is, we estimate that the contribution of so-

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

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

ciety at large in the socialization process of minorities is twice as large as the contribution implied by their actual representation in the population under random matching. This bias parameter is an indirect measure of immigrants' geographical and social segregation, e.g., in cultural-ethnic enclaves.

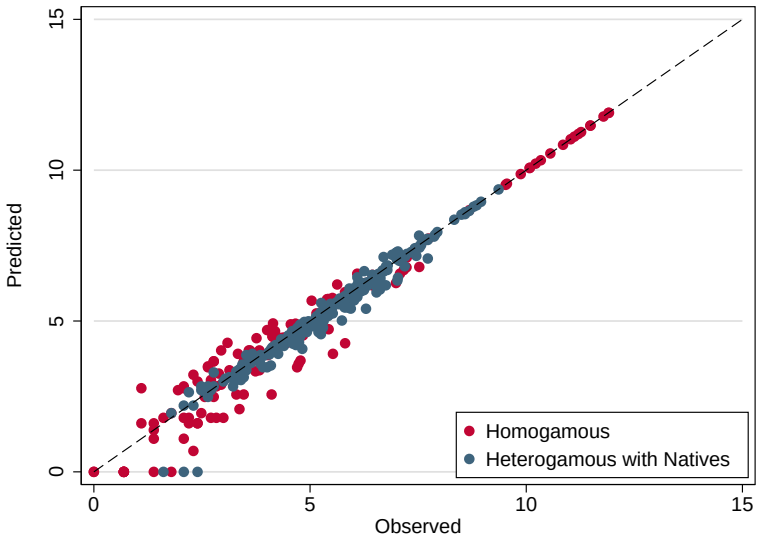
5.3 Validation and robustness

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

5.3.1 Validation

The first validation exercise consists in exploiting our preference parameters to predict the distribution of marriages observed for (out-of-sample) newly formed marriages from 2013 to 2019. The model fits out-of-sample marriage data very well; Figure 6 shows the relationship between the number of marriages observed in the years from 2013 to 2019 and the number of marriages predicted by our model by match and region. Figure C.9 displays the distribution of men and women along cultural lines in the sample and out-of-sample period.

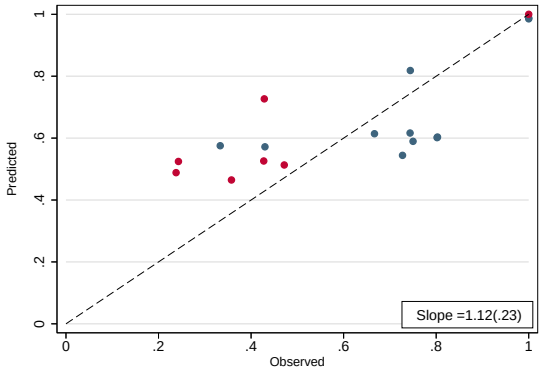
Figure 6: Model Validation - Non Targeted 2013-2019 Marriages by Match and Region



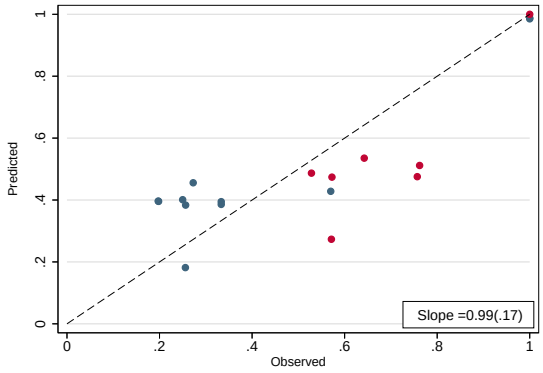
Notes: This figure displays the scatterplot of the relationship between the number of marriages observed in the out-of-sample data (in log) in the years from 2013 to 2019 and the number of marriages predicted by the model (in log) by region for homogamous families (red) and heterogamous families with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

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

(a) Italian Socialization



(b) Mother Socialization



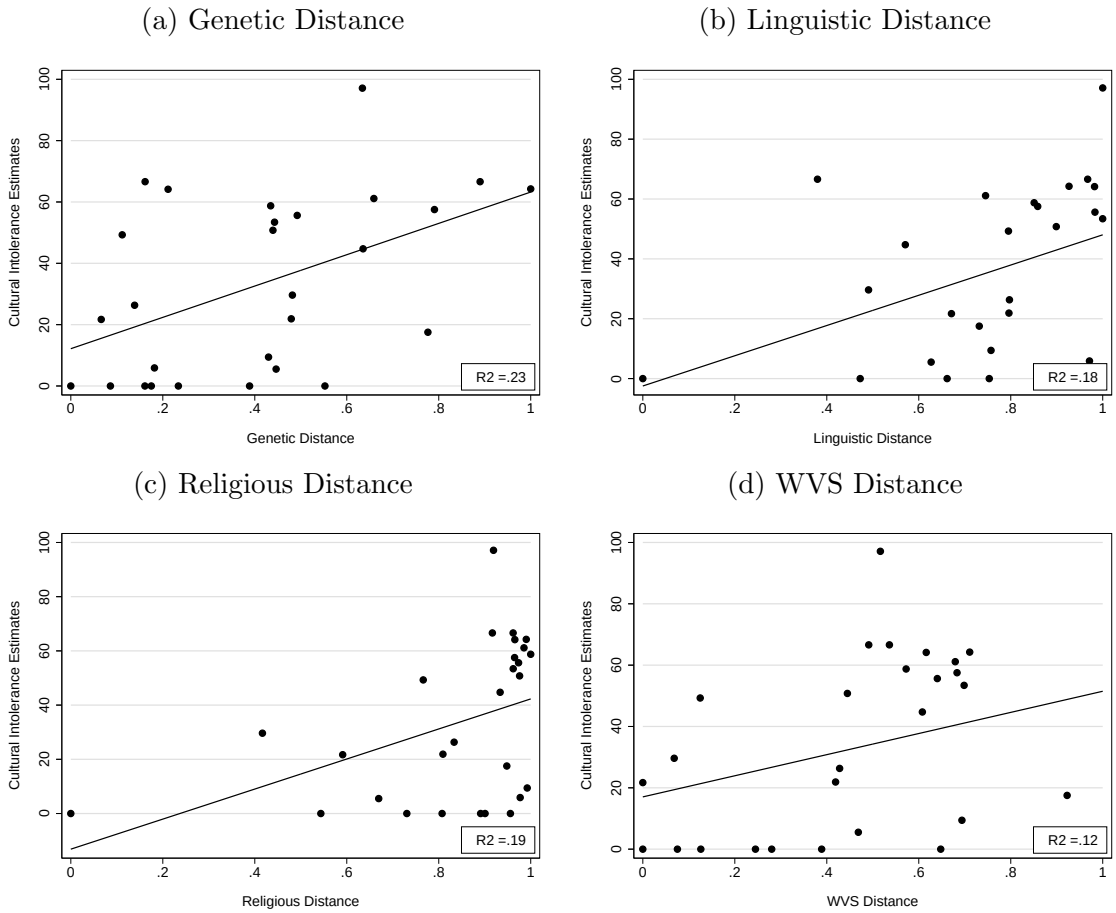
Notes: This figure displays the scatterplot of the relationship between the observed and predicted Italian and mother socialization probabilities for the subsample of marriages ending in divorce, separately for homogamous families (red), and heterogamous families with natives (blue).

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

Thirdly, we validate our estimates by asking whether they predict well the distribution of marriages by province, a higher level of geographical disaggregation; see Figure C.10. We exploit variation in the distribution of cultural-ethnic traits across provinces within the same region Indeed, we observe variation in the distribution of cultural-ethnic traits across provinces within the same region (within region heterogeneity accounts for about 1/3 of the total observed variation) and we show that our estimates, based on regional moments, in fact predict successfully the pattern of marriages at this more granular level.

Our last, but not least, exercise is focused on our main parameters of interest. We externally validate our cultural intolerance estimates (a revealed measure of perceived cultural distance) across cultural-ethnic groups by relating them to various cultural distance measures commonly used in the literature (Spolaore and Wacziarg, 2009, 2016), and discussed in Section 2. Figure 8 documents a systematic positive correlation between our cultural intolerance estimates and various cultural distance measures. These cultural measures explain from 15 to 24 percent of the variation in cultural intolerance. We find it remarkable as cultural distance measures are, by construction, symmetric, while our estimates do not impose any restriction on symmetry. Weighted estimates by the number of marriages per match type report a systematically higher correlation.

Figure 8: Cultural Intolerance Estimates and Cultural Distance Measures



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

5.3.2 Robustness

We discuss our main assumptions in turn.

Cultural-ethnic socialization is proxied with language socialization. Ethnic identity and spoken language are relevant culturally related specific attributes and both allow the direct transmission of cultural characteristics across generations.[45] We measure socialization by the *language spoken at home* within the family, as a form of parental cultural investment. While, reasonably, all children living in Italy learn Italian at school, speaking Italian at home when at least one spouse is an immigrant, in our interpretation, reveals deep-seated preferences for integration relative to ethnic identity (Bazzi et al., 2019; Salari, 2020). To corroborate

[45]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.

this interpretation, we provide evidence along two directions. First, we document that our measure of Italian linguistic socialization influences the achievement and educational choices of immigrant students. We obtained student-level data on standardized test scores in reading and math administered by INVALSI to all students in Italy at the end of grade 5. The test is identical for all students in a given grade, and it is blindly scored, hence results are fully comparable across schools. Crucially for us, INVALSI data also contain survey information on the main language spoken at home by students, as well as rich baseline information on students and family background. We describe in details the Italian educational setting and INVALSI data in Appendix B. Table 7 shows that speaking Italian at home improves the achievement in test scores of immigrant students both in reading and math. Immigrant students who speak Italian at home exhibit higher achievement by a 0.20 (0.11) standard deviations in reading (math). The point estimates are significant and relevant in magnitude. Moreover, by exploiting the longitudinal structure of the data, we investigate the language long-term impact on students' educational careers. We show that speaking Italian at home drives students into high demanding schools, i.e., Italian socialization at home (during the elementary school period) increases the probability of attending an academic or technical high school (as opposed to vocational one) by 2.7 percentage points, on a baseline rate of 82 percent for immigrant students in grade 10. This might ultimately have long-term implications for access to college and occupational careers (Brunello and Checchi, 2007; Carlana et al., 2021). Results are robust to various checks, see Appendix Table C.7.

Second, we provide survey evidence that Italian use at home is associated with weaker ethnic identity and stronger social integration networks, educational achievement and aspirations, and proficiency in the Italian language. More specifically, Table C.8 shows that children who speak Italian at home are: (i) about 50% more likely to have Italian friends (out-

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)
Province & Cohort FE	Yes	Yes	Yes
Student Controls	Yes	Yes	Yes
Family Controls	Yes	Yes	Yes
Observations	330,739	345,980	90,656
R-squared	0.144	0.098	0.058

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

side school), and are more likely to speak Italian with school mates and friends (columns 1-3); (ii) 15% more likely to achieve high education, to pass the academic year and have greater aspirations for their future educational career path (columns 4-6); (iii) significantly more proficient in Italian in writing, reading, speaking, and comprehension (columns 7-11).[46]

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

Marital matching is along cultural-ethnic lines only. In particular, we do not take into account sorting along other, in principle, important observable attributes, like education and age. This could potentially bias our estimates, if sorting along cultural lines of minorities were systematically associated with sorting along other observables. However, the introduction of flexible parameters representing the outside option of remaining single allows us to capture the multidimensionality of the marital selection process in a reduced form way. By representing a residual component of preferences, outside options capture differential sorting in both observable and unobservable attributes across cultural-ethnic groups and between homogamous and heterogamous families. Furthermore, in a related analysis, Adda et al. (2020) document strong preferences for cultural similarity relatively to other factors in a multidimensional model of marriage on Italian data, explicitly allowing for trade-offs between cultural distance, legal status, and other socio-economic spousal characteristics like education and wealth.[47]

Differences in fertility might be due to systematic variation in observables across households. To partial out these differences, we regress fertility on a set of covariates including marital duration, age at marriage of spouses, educational attainment, employment and occupational standing.[48] We then compute the fertility moments using the residual variability

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

[47]Specifically, Adda et al. (2020) estimate a reduction in marriage surplus for cultural heterogamy ranging from 7 percent (in intermarriages with European immigrants) to 25 percent on average (in intermarriages with Asian immigrants); while the reduction in marriage surplus for highly educated individuals in marrying low-educated spouses is just 7 percent on average.

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

out of this linear regression model, thus unrelatedly to potential confounding in our estimation. Figure C.11 reports cultural intolerances in line with our main estimates, while only the demand of cultural identity of East Asian minorities appears underestimated at baseline.

The distribution of the population across regions is exogenous. We abstain from modeling endogenous moving and/or residential location decisions. Endogenous moving or location decisions would be problematic for our estimates, if these decisions were motivated in part by marriage and socialization, as well as by unobserved heterogeneity. Consider the natural hypothesis that minorities that are particularly attached to their cultural identity choose to locate into more segregated areas. In this case, we would expect a positive correlation between vertical and horizontal socialization. Figure C.12 describes this relationship by plotting the probability of direct socialization in homogamous families over the corresponding population share by regions, showing instead a negative and, at times, statistically significant relationship. The evidence is consistent with direct cultural transmission within the family substituting horizontal socialization, i.e., if anything, minorities in more segregated regions display lower socialization rates to the language of parents. This substitution pattern, thus, represents a lower bound, net of complementarity in residential selection. To further alleviate potential concerns of endogenous location choices, we rely on pre-determined settlements of ethnic groups of immigrants across regions. Specifically, we predict the population distribution by ethnic group and region, by exploiting pre-existing variation in the geographical distribution of immigrants by ethnic group observed in 1993, interacted with subsequent inflows by origin, as in a shift-share instrument strategy (Card, 2001; Tabellini, 2020).[49] Indeed, the distribution of immigrants exhibit a strong geographical clustering along ethnic lines, and settlement patterns of ethnic minorities are a good predictor of location choices of newly arriving immigrants. Results in Table C.9 are in line with our baseline estimates.

6 Long-run integration patterns

In this section, we simulate the dynamics of the distribution of cultural-ethnic traits in the population induced by our structural model of marital matching, fertility, divorce, and socialization. While the exercise rests on the strong assumption that parameters are invariant over time, these simulations should be interpreted to highlight the implications of our estimated model with respect to the prospective pattern of cultural-ethnic integration

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

of different minorities in Italy.[50] It should also be noted that the notion of integration we are necessarily bound to adopt, given our data, refers to the practice of speaking Italian at home; that is, an individual belonging to a minority is integrated in our simulations when living in a household speaking Italian at home.

6.1 Population dynamics

The time unit in the simulations is a generation, i.e., a time interval of about 25-30 years. We fix the initial condition, generation $t = 0$, to coincide with the distribution of the population by region and ethnic group in our data. More precisely, while we observe the demographic characteristics of the Italian population over time, from 1995 to 2012, we interpret them for these simulation exercises as representing a cross section of the population in 2012, by region and ethnic group. Let this distribution be denoted p_t. The structural model we have estimated induces a map from p_t into p_{t+1}. Indeed, the model maps any distribution p_t into a vector of demographic characteristics of the population at time t, in terms of marital matching, fertility, divorce, and socialization, by ethnic group and region. The mapped fertility and socialization at t, by region and ethnic group, induces in turn a distribution of the population of the children of the population at time t, p_{t+1}.[51] The same simulation procedure, recursively, induces $p_{t+2}, p_{t+3}, ...$ [52]

The simulated long-run dynamics of the fraction of the population with cultural-ethnic trait i for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$ are reported in Figure 9, normalized so that $q_t^i = 1$ in $t = 0$ for comparability.[53] See Figure C.13 for non-normalized long-run dynamics of cultural traits. Despite estimates of cultural intolerance highlight immigrants' strong preferences for

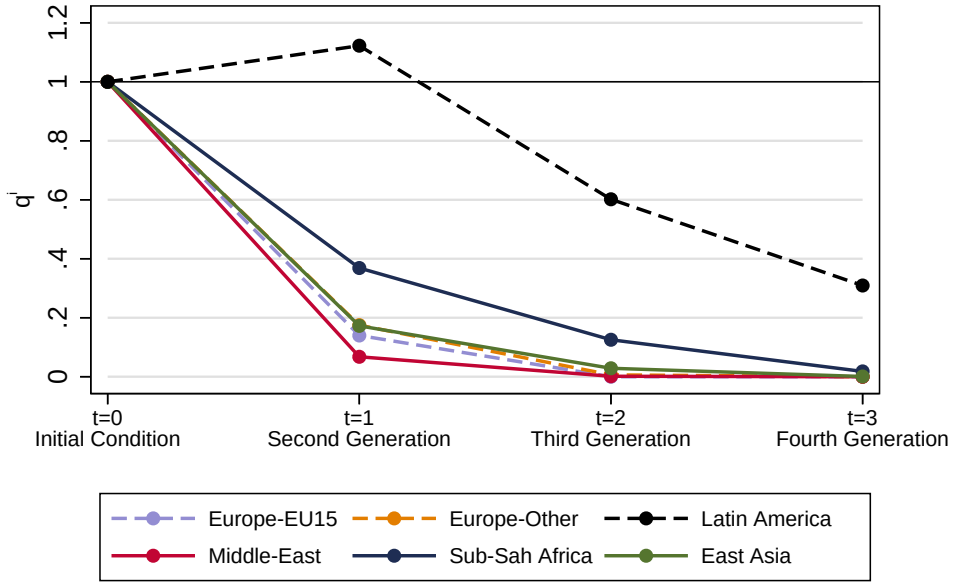
[50]The model might be generalized to allow for inter-temporal dependence of cultural intolerances, relaxing the assumption that cultural intolerance parameters are constant across generations. For instance, the cultural intolerance of second generations could be set equal to $\eta \Delta V_n^i$, where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

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

[52]At each step, we compute the marital matching equilibrium in the marriage market, represented by equation (4) subject to feasibility constraints in (2). This amounts to solving a system of $2K$ quadratic equations in as many unknowns, with K the number of cultural traits in the population, for each of the R regions. To this end, we take advantage of an iterative projection fitting procedure (IPFP) designed to find projections on intersecting sets of constraints, by projecting iteratively on each constraint (Galichon and Salanié, 2021; Galichon, 2018). For computational simplicity, the simulations assume $\alpha = 0$.

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

Figure 9: Long-run Dynamics of Cultural Traits (index=1 in $t = 0$)



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

maintaining their cultural identity, all cultural-ethnic minorities are simulated to integrate into the Italian majority along the language dimension. Indeed, the *integration rate*, defined as the reduction in the fraction of the total population (immigrants and natives) which is composed of immigrants who are not integrated to the native Italian culture, is 75% over the period of a single generation.[54] The rate of integration is however heterogeneous across cultural-ethnic groups. In particular, the European-EU15 and Other European minorities integrate almost completely, 87%, in a single generation. A similar pattern is also displayed by the North African-Middle Eastern minority. On the other hand, a significantly slower integration rate is achieved by the Latin American minority, whose generation at $t = 1$ is even less integrated than the generation of their parents in the simulation, and reaches an integration rate of 70% only by the fourth generation. A slower integration rate also characterizes the East Asian and Sub-Saharan African minorities, 82% and 63% in one generation. For third generations overall 93% of immigrants converge towards Italian culture in about 50-60 years.

The patterns of cultural integration of European-EU15 and Other European minorities

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

are the result of their relatively low cultural intolerance preferences. In a similar way, the East Asian and Sub-Saharan African minorities' slower integration is due in part to their higher intolerance parameters. But intolerance parameters are not the only determinants of the dynamics of integration of different cultural-ethnic groups. Homogamous marriage rates, fertility rates, and other demographic characteristics in fact turn out to have sizable independent effects on cultural integration in the simulations.[55] This is clearly illustrated by the fact that, while North Africa-Middle East, Sub-Saharan Africa, and East Asia show relatively comparable cultural intolerance preferences, they display significant differences in the dynamics of integration. Indeed, a strong estimated selection into homogamous marriages of immigrant from Sub-Saharan Africa allows them to sustain their cultural heterogeneity by accessing superior direct socialization technologies; see Figure C.14 for evidence on the evolution of the homogamous marriage (panel a) and intermarriage (panel b) rates over successive generations. On the other hand, estimated fertility rates are particularly high for East Asian minorities and this is a fundamental factor behind this minority's integration pattern. Finally, the relative success of Latin America in securing their cultural distinctiveness over time is due in large part to the fact that they turn out to be uniquely able to socialize children also in heterogamous marriages with natives.

6.2 Counterfactual cultural intolerance parameters

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

We consider, first, the case in which $\Delta V_n^i = 0$ for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$; that is, we consider a counterfactual environment in which natives are fully tolerant towards minorities, and offer complete acceptance of the immigrants' cultural diversity. Results, displayed in Figure 10, show that this counterfactual experiment induces on average a 15% increase in the fraction of the population composed of non-integrated immigrants after one generation (the integration rate, as we have defined it, is negative). In the long-run, immigrants start integrating but slowly, so that the heterogeneity in the cultural traits in the population shows a remarkable persistence. While immigrants maintain strong preferences for cultural identity, an increase in the acceptance of natives has the effect of making intermarriages of

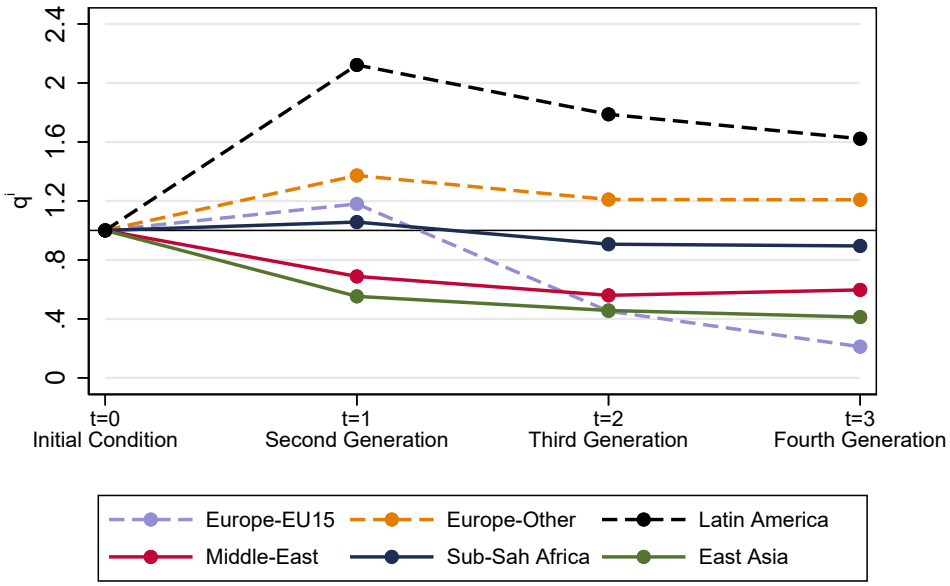
[55] With respect to fertility, this is the case even though predicted fertility rates for all groups are below reproduction level, which potentially has implications on marriage market competition as well.

immigrants with natives more valuable for both. Indeed, the coordination failure in children socialization that in the general framework undermine the value of intermarriages is now muted, and the immigrant spouse is able to socialize children to his/her own trait. In our counterfactual simulation, we observe a large increase in intermarriages with natives, and in parallel a lower demand for homogamous marriages. Furthermore, fertility rates in intermarriages with natives are substantially increased. Socialization to the Italian language is reduced, as it is driven only by horizontal socialization of society at large. All these effects induce a reduction in immigrants' integration compared to the baseline; see Figure C.15 and Figure C.16 for a representation of the mechanisms driving the dynamics of integration.

The dynamics of integration is still heterogeneous however, across groups. Specifically, we find that, by the fourth generation, the integration rate of Europe-EU15, East Asia and North Africa-Middle East is about 80%, 59% and 40%, respectively. Instead, Other Europe and Sub-Saharan Africa display a much lower integration rate. Once again, the Latin American minority appears to be an outlier, as its fraction in the population increases by 50% by the fourth generation.

We study also the extreme opposite case in which Italians are fully intolerant towards all minorities, $\Delta V_n^i = 100$ for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$. The dynamics of integration of ethnic

Figure 10: Long-run Dynamics of Cultural Traits with Italians Fully Tolerant, $\Delta V_n^i = 0$ (index=1 in $t = 0$)

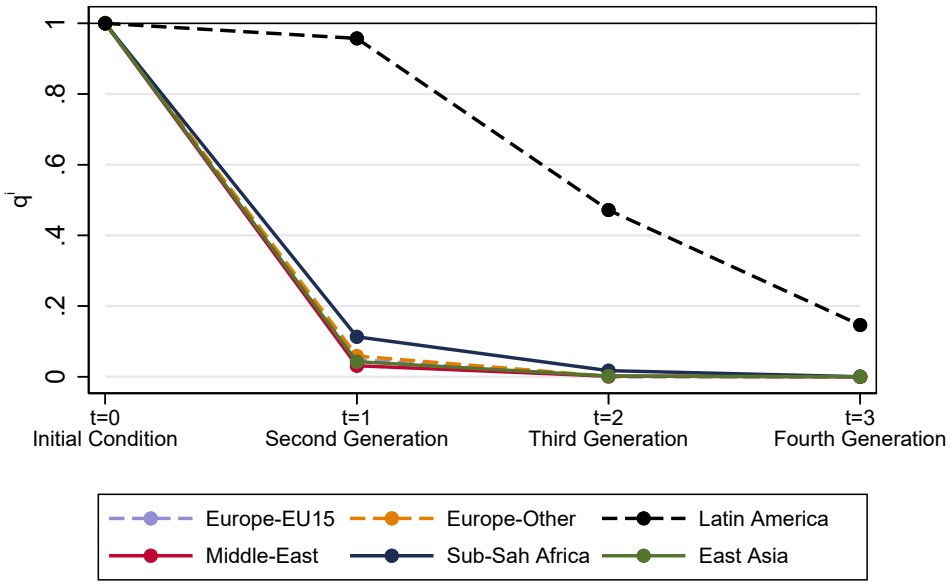


Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations assuming the case of complete tolerance of Italian majority towards minorities.

minorities follow the results in the baseline, with an integration rate of about 80% overall over a single generation; see Figure C.17. The most significant difference in integration is attributable to Latin Americans, whose integration rate is now (positive and) as high as 47% in a single generation.

These results are in line with those obtained by Fouka (2020), who examines the immigrants' response to the shift in attitudes towards immigration in the US in the aftermath of WWI, which lead to imposition of language restrictions in elementary schools. She shows that individuals subject to language bans i) were less likely to volunteer in WWII, ii) were more likely to marry within their ethnic group, and iii) strengthen the vertical socialization of their children (high foreign names).[56] In our model, this shift in attitudes maps into an increase in ΔV_n^i (lower supply of acceptance on the part of natives), while the language ban maps into higher costs of socialization. According to our counterfactual analysis, this would induce higher homogamy in marriages and higher socialization rates to the culture of minorities, in line with Fouka (2020)'s results.

Figure 11: Long-run Dynamics of Cultural Traits net of Economic Integration Incentives, proportional increase in ΔV_i^n (index=1 in $t = 0$)



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, given an increase in ΔV_i^n for all minorities by 20% of ΔV_n^i.

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

To study the potential role of economic incentives for integration, we consider the case in which ΔV_i^n increases for all $i \in \{i_E, i_O, i_M, i_A, i_S, i_L\}$ by 20% of ΔV_n^i, which could reveal either a reduction in economic incentives to integration (e.g., the job opportunities available for the more integrated immigrants) or stronger cultural intolerances. When we simulate our model in this case, the dynamics of integration of minorities towards Italian culture accelerate by 10 percentage points compared to baseline estimates.Results, displayed in Figure 11, appear counterintuitive. In principle, we might have expected that, in the absence of economic incentives, the strengthening of migrants' cultural preferences would make their cultural integration more demanding and hence slower, by providing higher socialization rates and reducing heterogamous marriages. On the contrary, we show that the reduction in economic incentives to integration contributes to accelerate cultural convergence.

Indeed, the stronger attachment of minorities to their identities makes marriage riskier and costlier, by commanding a larger investment of resources in socialization. The value of marriage significantly reduces for both homogamous marriages of immigrants and heterogamous marriages (while, the value of homogamous marriages of natives remains unchanged), leading to a general equilibrium effect of compression of the marriage market for immigrants. Hence, cultural convergence in this setting is induced by a lower participation in the marriage market and lower fertility of immigrants. In other words, in this counterfactual, the acceleration of integration is mostly an effect of the reduction in the population growth for immigrants with respect to natives. The probability that a child with an immigrant parent is integrated to the Italian culture is lower in the absence of economic incentives (i.e., the parental socialization effort becomes more effective), but the fraction of the total population (immigrants and natives) belonging to the native culture, on the contrary, is higher.[57]

These results are in accordance with those of Adda et al. (2020), who study how marriage and separation choices of immigrants respond to a change in the labor market value of marriage. Specifically, Adda et al. (2020) study the effects of the EU enlargements in 2004 and 2007 which,[58] by guaranteeing permanent legal status and the right to work to the citizens of the new Eastern European member-states, eliminate their prior benefits of marrying an EU native. This paper shows that, after these enlargements i) the probability of the relevant intermarriages decreases, and ii) the hazard rate of separation of these intermarriages increases. In our model, this experimental setting maps into a reduction in economic incentives to integration, a larger ΔV_i^n, which translates into the effects on intermarriages and separations observed by Adda et al. (2020) after the enlargements.

[57] Other counterfactuals changing cultural intolerance parameters to simulate demand and supply effects in the dynamics of integration at equilibrium are discussed in the Appendix; see Figure C.18.

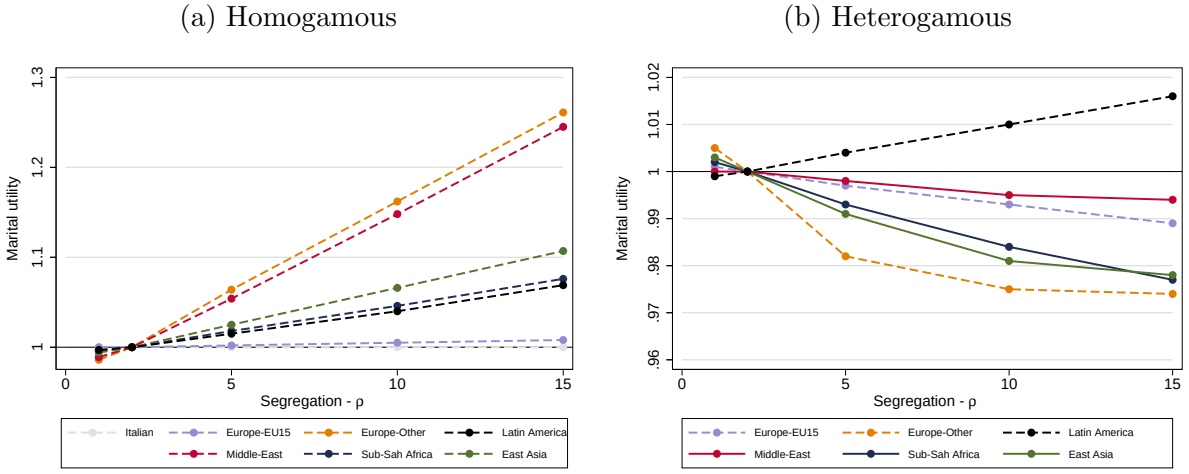
[58] For further details, we refer to footnote 8.

7 Policy and welfare

Cultural integration of immigrants in our structural model of marriage, fertility, and socialization is the equilibrium outcome of two interacting mechanisms, frictionless matching and collective decisions at the household level. Both these mechanisms are efficient.[59] A *dynamic* welfare analysis of equilibrium in our model however displays an externality in the fertility and socialization choices of households. Individual households at time t in fact do not take into account the aggregate effects of their choices on the size and the composition of the population by cultural-ethnic group in the future, from time $t+1$ onwards. A dynamic welfare analysis of policy interventions in our model can be performed with a utilitarian social welfare function W, which aggregates the utilities of all (groups of) agents in the economy, in different family types and ethnic groups, over multiple periods. We report next on such welfare analysis for a specific policy choice, the social segregation of immigrant minorities.

Consider the parameter ρ in our model, which we introduced to represent the strength of the immigrants' cultural-ethnic network through a segregation bias in their pool of reference. We consider ρ as, at least in part, a policy variable. In fact, geographic and social segregation of immigrants (e.g., in living quarters, schools, religious gathering, social clubs, etc.) is the result of their own choices as well as of governmental and local policies regarding e.g., public

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



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

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

housing, freedom of religion, schooling, local police, etc. We have estimated ρ of about 2 in our empirical analysis. We simulate now the effects of policies which change ρ to a range of values, from 1 to 15, at time t, before agents match in the marriage market. While we solve for the whole equilibrium at t and $t+1$, we concentrate here on the welfare effects of the segregation bias ρ. At time t, as ρ increases, immigrants benefit from a more efficient socialization technology (via the horizontal channel), and hence homogamous marriages increase largely in value (i.e., expected marital utility) and in number. Instead by exacerbating the socialization conflict, social segregation decreases the value and hence the number of heterogamous marriages with natives.[60] Figure 12 reports the marital utility response to an increase in ρ for homogamous (panel a) and heterogamous families (panel b). Finally, homogamous marriages of natives also increase in number via market equilibrium, even though they are not affected in value as they socialize their children at no cost; that is, independently of ρ.

Panel a. of Table 8 reports the number of marriages and the utilitarian welfare com-

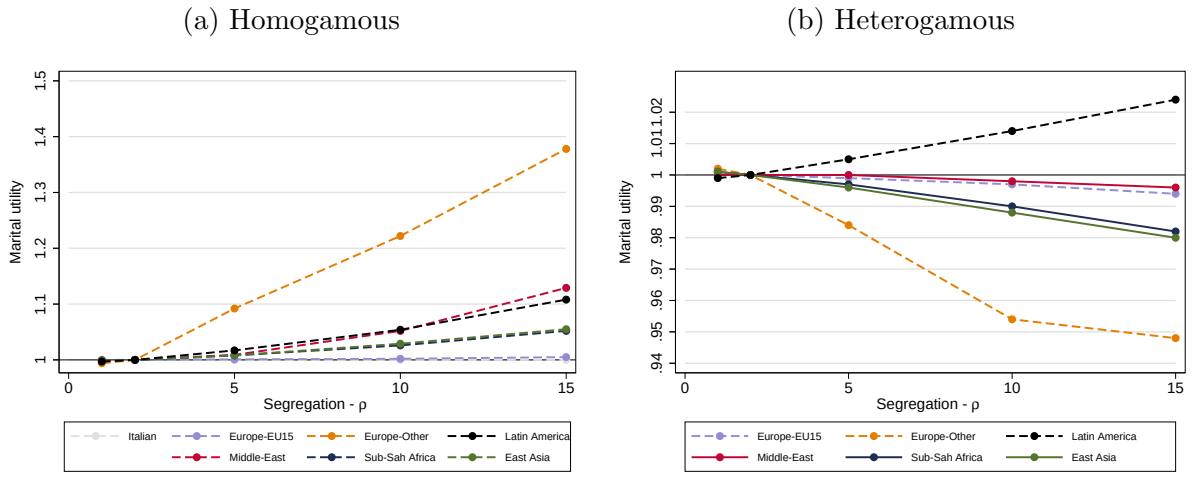
Table 8: Social Welfare Response to Segregation by Family Type and Ethnic Group

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

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

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

Figure 13: Marital Utility Response to Segregation at $t+1$, by Family Type



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

puted at time t, by family type (columns 1-6). Social welfare increases monotonically with segregation bias ρ for homogamous marriages of both immigrants and natives, while the heterogamous marriages display a welfare loss. We also report welfare results at the individuals level distinguishing natives from immigrants, i.e., lumping immigrants of all groups into a single average category (columns 7-8 of Table 8). Under natural assumptions,[61] segregation ρ proves to be welfare improving for any weight structure, as both natives and immigrants realize welfare gains.

The most interesting effects appear at time $t+1$, when the externality in the fertility and socialization choices of households plays a fundamental role. As represented in Figure 13, the marital utility of homogamous (panel a) and heterogamous families (panel b) calculated at time $t+1$ respond to segregation in a similar direction as at time t. But, the fertility and the socialization rate in homogamous immigrants households at time t grows steeply with ρ affecting the size and composition by cultural-ethnic group of the population at $t+1$. As ρ increases then, the larger fraction of second-generation immigrants available in the marriage market increases the number of both homogamous immigrant and heterogamous marriages, in turn increasing their contribution to the social welfare function. Panel b. of Table 8 reports the marriage and welfare effects computed at time $t+1$. Maintaining the

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

assumption that utility is transferred proportionally to the spouses outside options, we have that social welfare increases with segregation bias ρ for any weight structure, as overall both natives and immigrants realize welfare gains. The increase in heterogamous marriages compensate in the aggregate the reduction in the contribution of homogamous marriages of natives (which have constant value but decrease in numbers) to the social welfare function.[62] In consideration of the demand of minorities to preserve their cultural identity, segregation policies lead to positive short-run and dynamic welfare effects.

8 Counterfactual migration inflows

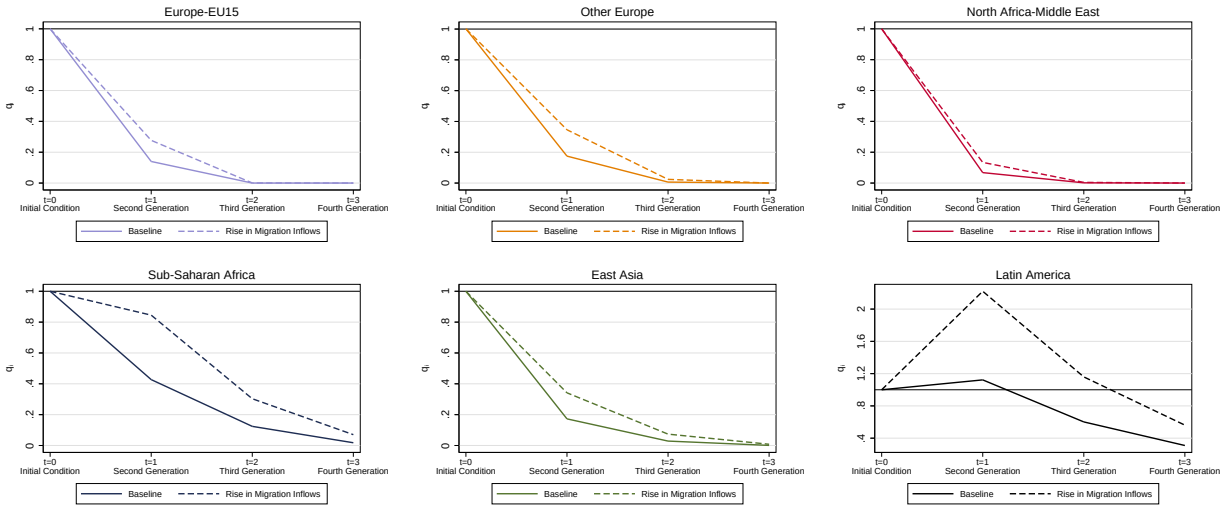
In the last few years, Italy has experienced a significant increase of migration inflows, mainly originating from Sub-Saharan Africa and Middle-East countries. We study the effects of such a rise in immigration on cultural heterogeneity in Italy, by performing two counterfactual exercises. In both cases,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

Keeping constant the share of each group, Figure 14 compares the dynamics of the distribution of cultural-ethnic traits in the population at the baseline (solid line) with the distribution resulting from the rise in immigration (dashed line). Overall, doubling the shares of second-generation minorities at $t = 1$, leads to a reduction in the integration rate of 7 percentage points for third generations, 86% compared to 93% at the baseline. More in detail, the rise in migration inflows has no effect on the cultural integration of European-EU15, Other European and North African-Middle East minorities already in the third generation. On the contrary, the incoming waves of Sub-Saharan Africa and East Asia immigrants produce an effect in delaying their full convergence to host country cultural practices. In particular, we estimate a 20 and 6 percentage points reduction in the integration rate of Sub-Saharan African and East Asian minorities, respectively, compared to baseline.

We observe qualitatively similar results (even though stronger in magnitude) in the second exercise, when we modify the relative distribution of second generations, overweighting North Africa-Middle East, Sub-Saharan Africa and East Asia; Figure 15 displays the resulting integration response. For a comparable increase in migration flows, the three groups

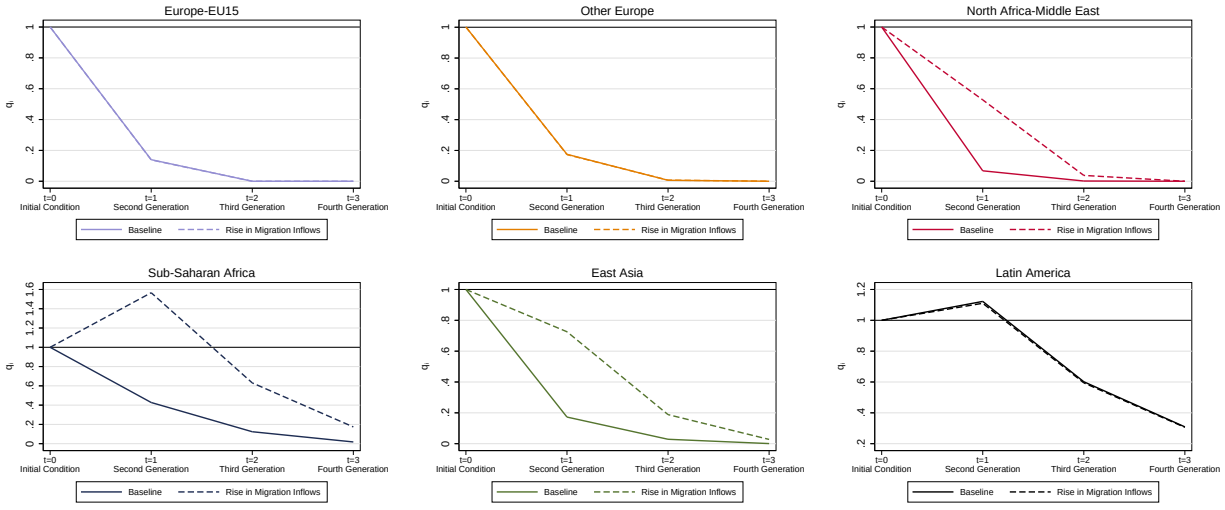
[62] The marriage market endogenously increases with ρ at time $t + 1$, while the number of singles reduces.

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



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations, normalized so that $q_t^i = 1$ in $t = 0$. The solid line represents the dynamics at baseline, while the dash line represents the dynamics after doubling the share of second generation immigrants, proportionally for all minority groups. Figure C.19 reports the non-normalized long-run dynamics.

Figure 15: Long-run Dynamics with Compositional Increase in Migration Inflows



Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for minority groups, over successive generations, normalized so that $q_t^i = 1$ in $t = 0$. The solid line represents the dynamics at baseline, while the dash line represents the dynamics after raising the share of North Africa-Middle East, Sub-Saharan Africa and East Asia second generation immigrants. Figure C.20 reports the non-normalized long-run dynamics.

highlight significant differences in integration patterns, with Sub-Saharan African and East Asian minorities accentuating their successful transmission of cultural values dramatically.

In particular, the cultural integration response of North Africa-Middle East immigrants to the exogenous rise in inflows is reduced by only a 4 percentage points. The response of East Asian and Sub-Saharan African minorities to a comparable variation ranges from 20 to 60 points, slowing down the process of cultural integration.

9 Conclusions

As cultural boundaries are increasingly salient, the design of adequate and successful policies to integrate minorities is a fundamental and challenging policy objective of modern societies. In this paper, we offered a new perspective to interpret cultural integration as an equilibrium outcome of marital matching and collective household decisions. We show by counterfactual analysis how the dynamics of immigrants integration over time respond to variations both in the demand of immigrants for the preservation of their cultural identity, as well as in the supply of acceptance of the immigrants' cultural diversity on the part of natives. These findings have in principle novel implications for the evaluation of different immigration policies, beyond across-the-board integration on one side and restrictive closed-border policies, on the other. Indeed, several results we obtain might not clearly emerge without an integrated equilibrium analysis of matching, fertility and socialization: i) an higher acceptance of the culture of minorities on the part of natives reduces the integration rate of immigrants; while ii) a reduction in economic incentives to immigrants instead accelerates their integration rate.

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

References

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

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

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

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

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

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

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

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

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

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

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

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy 81*(4), 813–846.

Becker, G. S. (1974). A Theory of Marriage: Part II. *Journal of Political Economy 82*(2), 11–26.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

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

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

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

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

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

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

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

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

For Online Publication

A Data and Sample Construction

We obtained restricted access administrative Italian data at the individual level from ISTAT through its ADELE Laboratory.[63] 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. We exploit marriage records from municipal vital statistics registries to recover matching patterns by ethnic group of the spouses. Marriage records contain the universe of marriages celebrated each year in Italy from 1995 to 2012. They provide information on the main socio-demographic characteristics of the spouses. They are collected through the ISTAT model compiled by the Registrar of the City Civil State in which the marriage took place. For each marriage, the section dedicated to the wedding reports: the date of marriage, the type of ceremony (religious or civil), the municipality of the ceremony and the choice of the property regime by the spouses (community or separation property). The information provided for each spouse includes: date of birth, municipality of birth, municipality of residence at the time of marriage, the place of future residence of the spouses, the previous marital status, the education level, the employment status, and for immigrant individuals the nationality and the country of origin. In order to account for 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.

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

[63] 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 analysis is made available upon inspection by ADELE officers in compliance with the laws on the protection of statistical confidentiality and of personal data. For further information, visit https://www.istat.it/it/informazioni-e-servizi/per-i-ricercatori/laboratorio-adele.

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

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

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

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

A.2 Dataset Construction and Empirical Moments

The empirical estimation is based on a unique dataset that links households information across different sources. We matched marriage, birth and separation records on the exact date of marriage and spouses' exact date and place of birth (Italian province for natives and country of origin for foreigners), which are reported in all registries. In the birth records matching, the combination of these characteristics allows for an exact one-to-one matching for 98.8% of marriages, while in the separation matching, we match exactly the 99.5% of marriages, and we discard the remaining fraction. Such low percentages suggest that marriages can be uniquely identified through the set of time-invariant characteristics listed above. The final sample of marriages (4,151,551) corresponds to 92.58% of the universe of marriages celebrated in Italy during the time interval 1995-2012. In the final dataset the fertility rate corresponds to 69.56% with an average of 1.54 children per family. Of all marriages, 7% end up in separation in the first years of the marital union.

From this final sample, we recover the following empirical moments. The marital utility net of the outside options of singlehood $\hat{U}_{hj}$ for the household of type hj is identified from equation (4), exploiting the number of hj marriages formed in each region r, μ_{hj}, and the number of unmatched men of type h and women of type j for each region r, $\mu_{h.}, \mu_{.j}$.

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

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

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

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

A.3 Language Socialization

Socialization data come from the *Condition and Social Integration of Foreign Nationals* Survey, conducted in 2011 and 2012 in all Italian regions on a sample of 9,600 families. The survey targeted foreign residents in Italy and it was conducted at the household level to provide socio-demographic information about all family members, for a total sample of 25,356 respondents. The aim of the survey was to collect essential aspects of the socio-economic integration process of immigrants in Italy, with a particular focus on linguistic integra-

tion. Different dimensions have been targeted such as: family composition, educational level, migratory path, employment status, discrimination and integration perception, living environment conditions, religious affiliation, social network formation and socio-political participation. 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. 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.[66] In each selected household, all members were interviewed, both foreign-born and natives.

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

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

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

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

[67] 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?*

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

B Education in Italy and INVALSI data

We obtained administrative student-level data on standardized reading and math proficiency tests, as well as related survey data from the National Institute for the Evaluation of the Italian Education System (INVALSI).[68] In what follows, we start describing the Italian educational system, our data sources and variables of interest; passing then onto a discussion of our sample construction.

B.1 Italian educational setting

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

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

B.2 INVALSI tests and survey data

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

[68] 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/.

mean and standard deviation of one. The test consists of multiple-choice and open-ended questions, where the exact structure varies by grade.[69]

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, grade retention, and family background characteristics such as parents' education, migration history, employment status, and some measures of socio-economic status. Crucially for us, INVALSI also collects additional data from a students' questionnaire, including a specific question on the main language spoken at home by students, similarly to our main analysis. The questionnaire is administered only in grade 5 (the last year of elementary school) and in grade 10 (second year of high school).

B.3 Sample selection

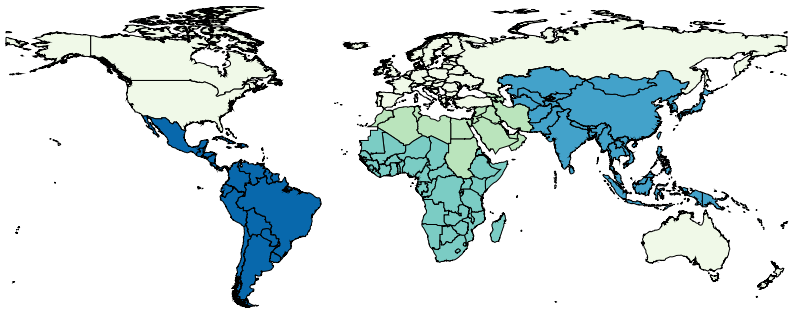
In our analysis, we focus on students enrolled in grade 5 between school years 2012-13 and 2018-19. For these pupils the language spoken at home is likely a choice of the parents. Thanks to a unique student identifier, we are able to follow students over time and match the scores and information of students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

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

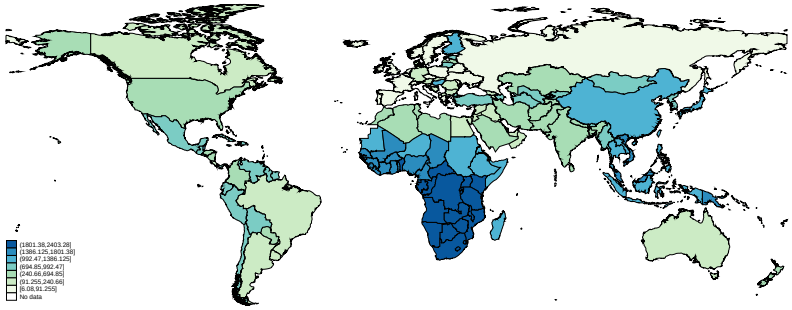
C Additional Figures and Tables

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

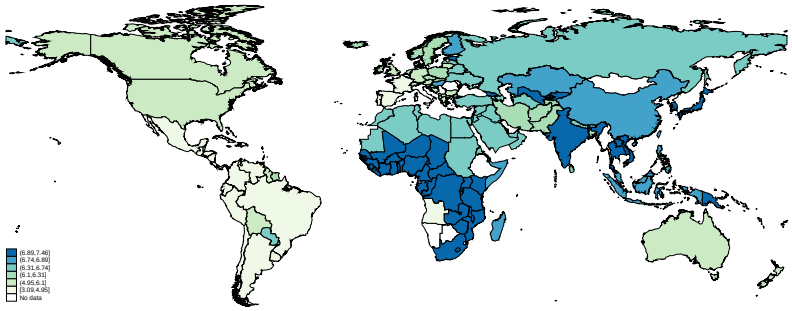
(a) Our Cultural-Ethnic Group Classification



(b) Genetic Distance with respect to Italy

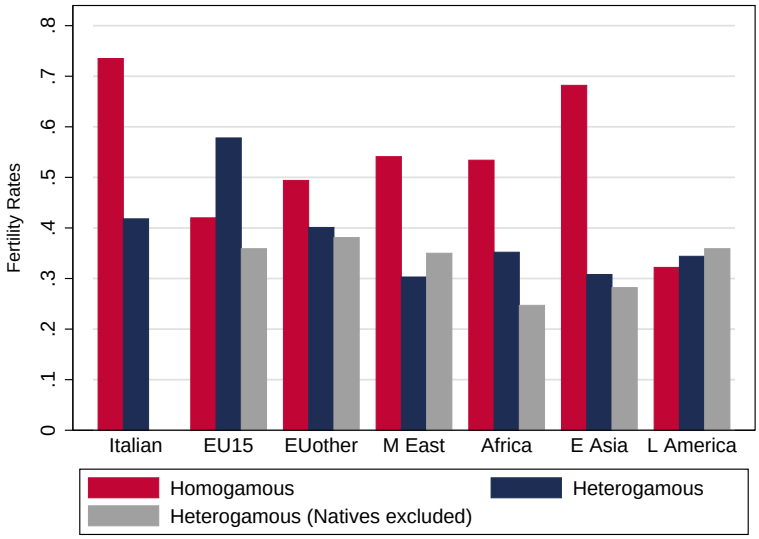


(c) Linguistic Distance with respect to Italy



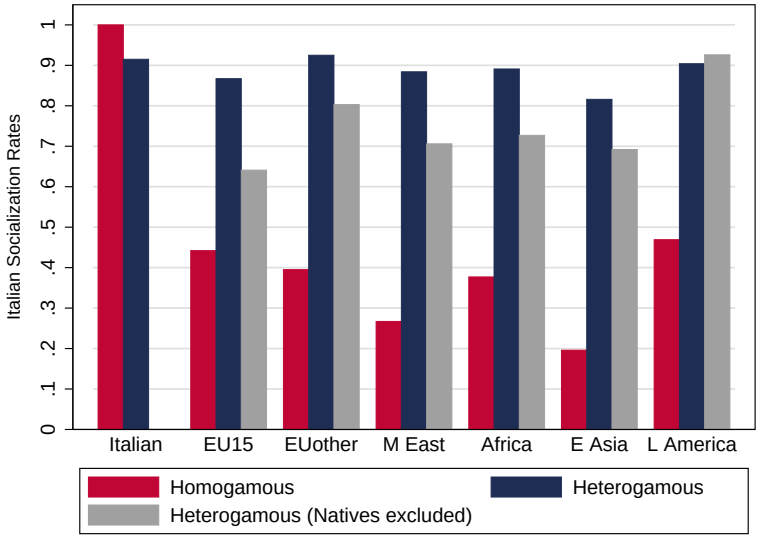
Notes: This figure shows our classification of countries in cultural-ethnic groups (panel a) and plots the cultural distance of each country towards Italy as proxied by genetic (panel b) and ethnolinguistic distance (panel c). Data for genetic and ethnolinguistic distance are available thanks to Spolaore and Wacziarg (2016).

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



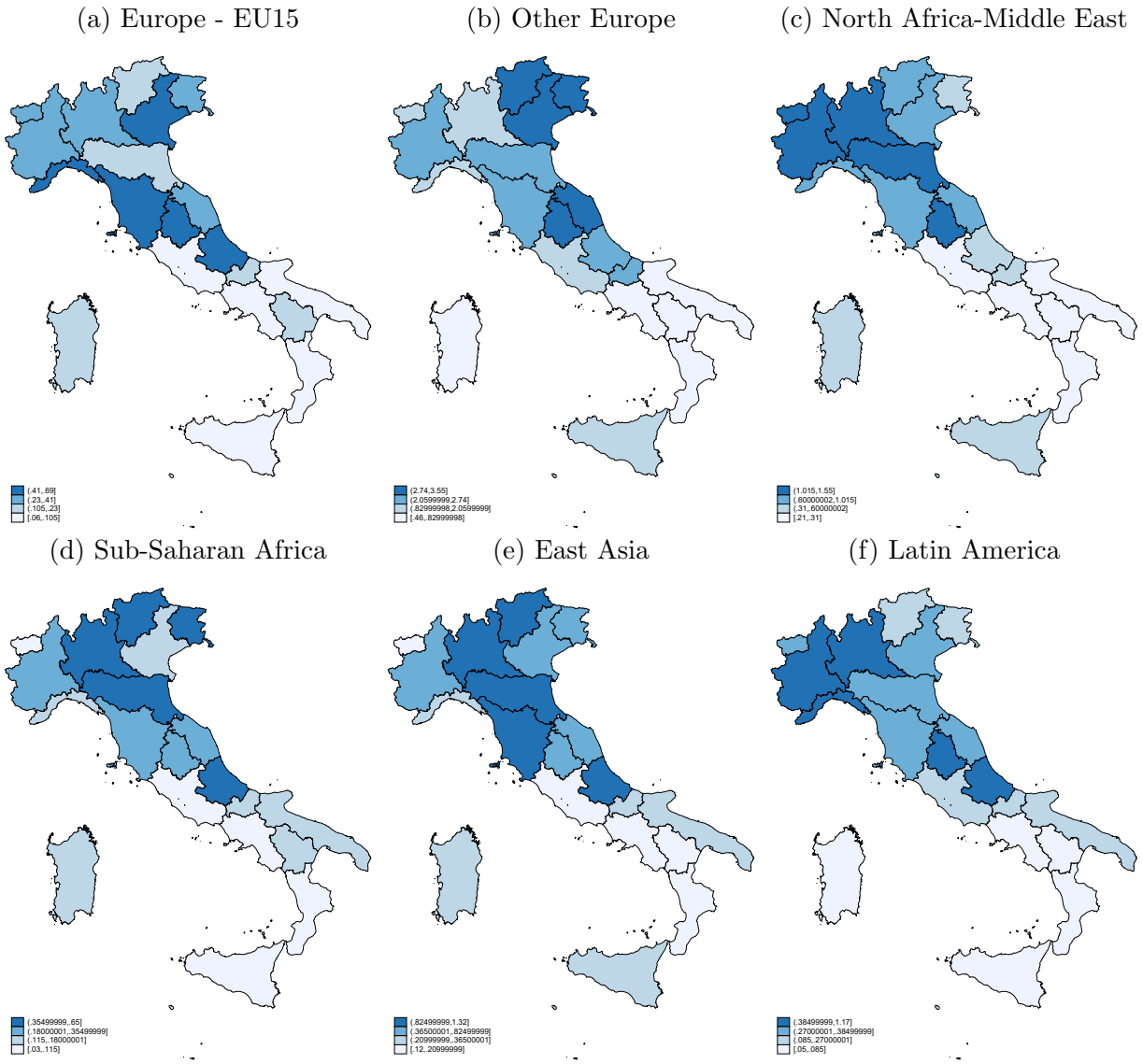
Notes: This figure shows fertility rates by ethnic group of spouses. Estimates are reported separately for homogamous, heterogamous, and heterogamous families excluding marriages with natives.

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



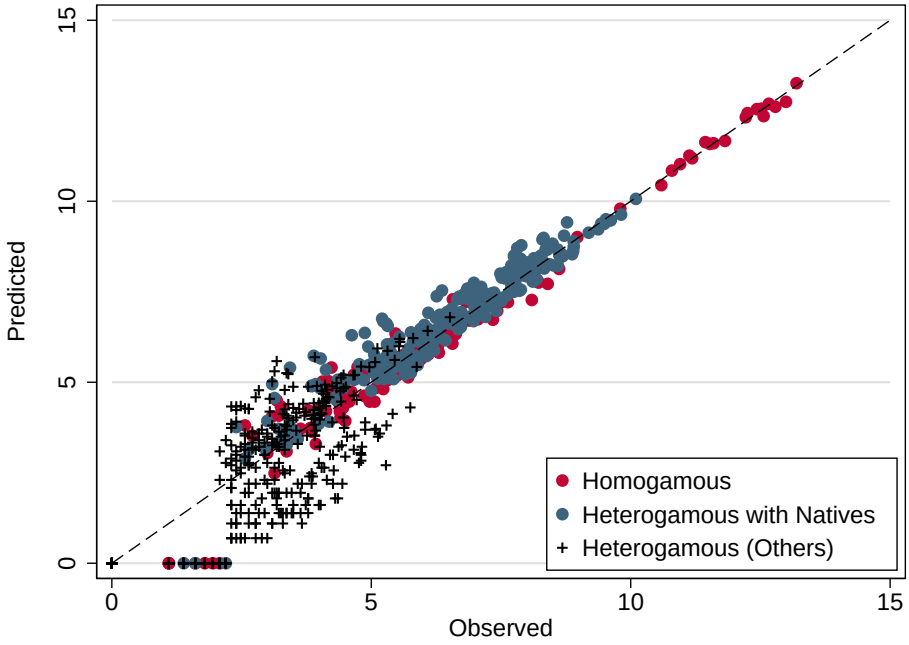
Notes: This figure shows Italian socialization probabilities by ethnic group of spouses. The outcome variable is an indicator for whether the child speaks Italian within the family. Estimates are reported separately for homogamous, heterogamous, and heterogamous families excluding natives.

Figure C.4: Migrants' Distribution across Regions



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

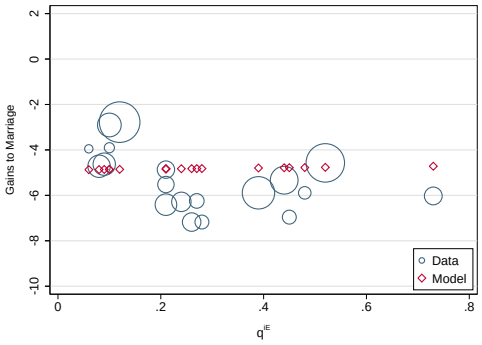
Figure C.5: Fit of the Model - Number of Marriages by Match and Region



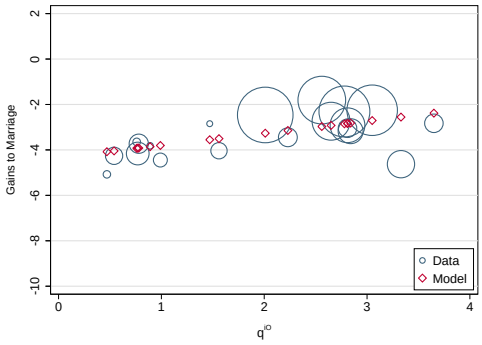
Notes: This figure shows the relationship between the number of marriages observed in the data (in log) and the number of marriages predicted by the model (in log) by region for homogamous families (red), heterogamous families with natives (blue), and all other heterogamous matches (black).

Figure C.6: Fit of the Model - Gains to Marriage for Homogamous Families by Region

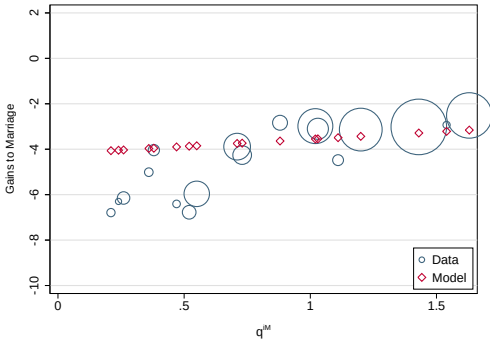
(a) Europe-EU15



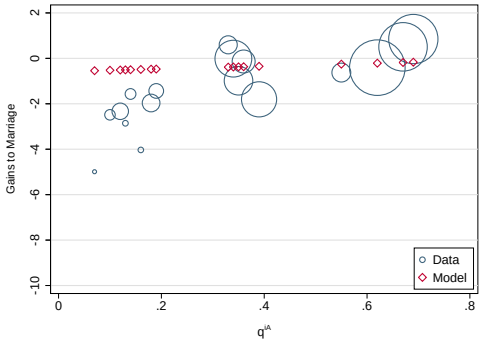
(b) Other Europe



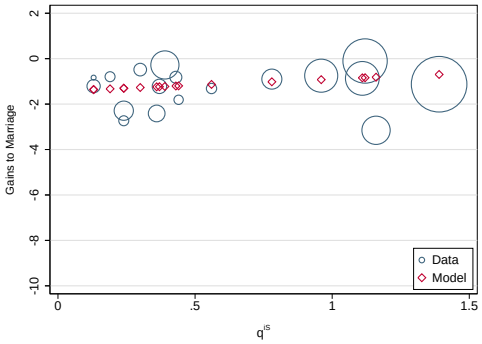
(c) North Africa-Middle East



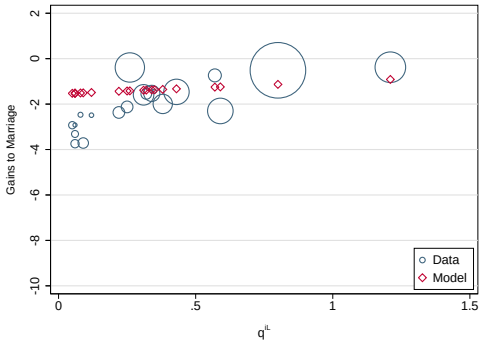
(d) Sub-Saharan Africa



(e) East Asia



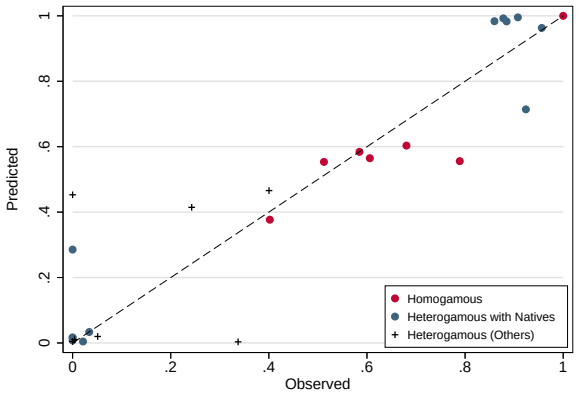
(f) Latin America



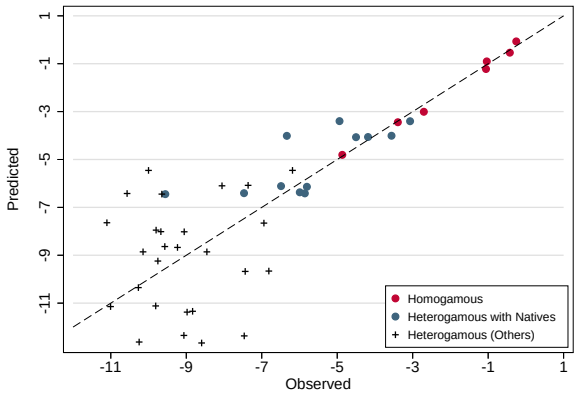
Notes: This figure shows predicted and implied gains from marriage for homogamous families of ethnic group minorities over the corresponding population share, q^i (in percentage), by region (average over the time period). Empirical moments are weighted by the observed number of marriages per region.

Figure C.7: Fit of the Model - Socialization Rates and Gains from Marriage

(a) Father Socialization



(b) Gains to Marriage

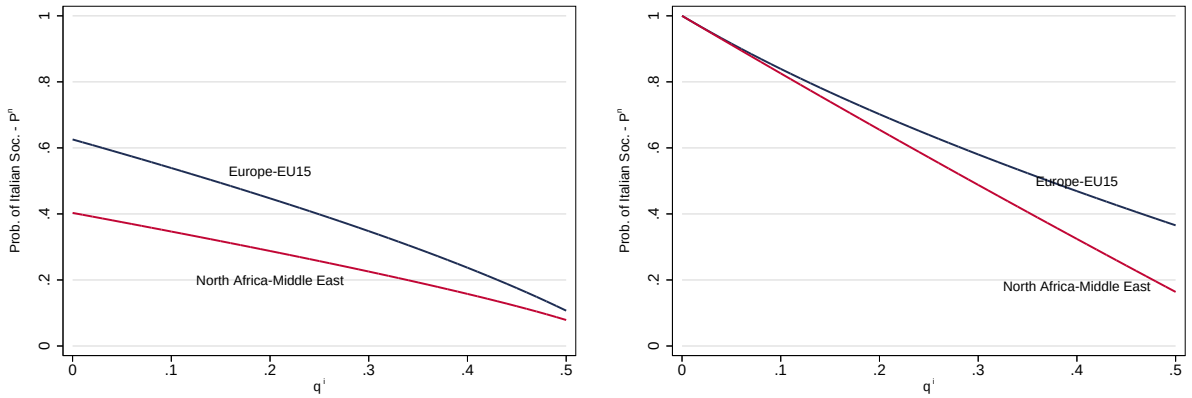


Notes: This figure shows the average fit of the model considering socialization probabilities (panel a) and gains to marriage (panel b), separately for homogamous families (red), heterogamous families with natives (blue), and all other heterogamous matches (black).

Figure C.8: Estimates of Italian Language Socialization by Minorities

(a) Homogamous

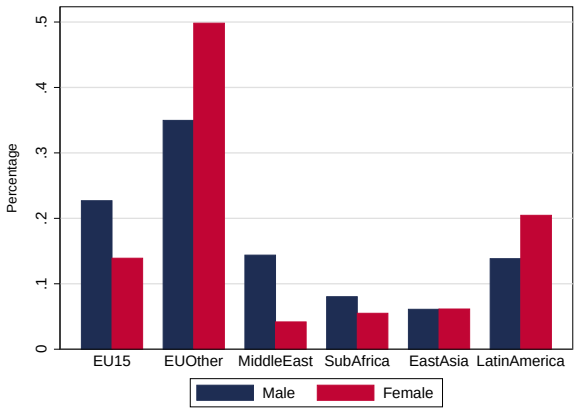
(b) Heterogamous



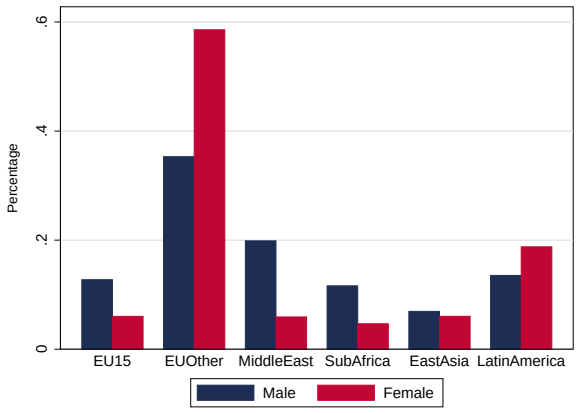
Notes: This figure reports estimates of the Italian socialization probability, P^n, of Europe-EU15 and North Africa-Middle East minorities over the potential population share, q^i. Estimates for homogamous families are in panel a, and estimates for heterogamous marriages with natives are in panel b.

Figure C.9: Distribution of men and women by ethnic group, in the sample and out-of-sample period

(a) Sample period (1995-2012)



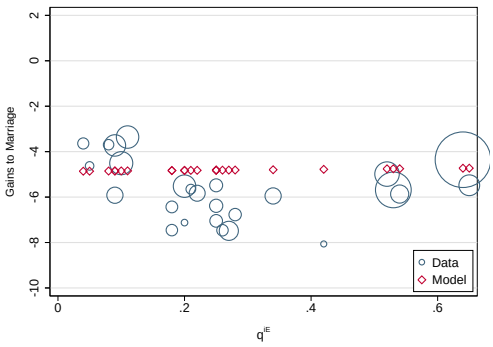
(b) Out-of-sample period (2013-2019)



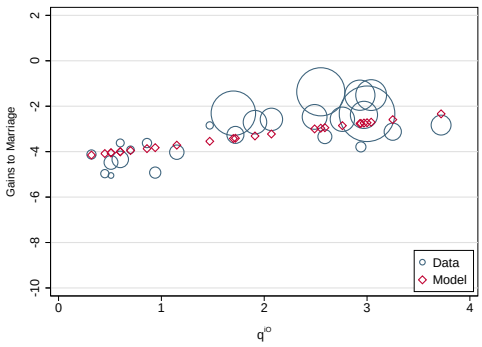
Notes: This figure shows the distribution of male and female population vectors (conditional on getting married) by cultural-ethnic group in the sample (1995-2012) and out-of-sample (2013-2019) period, in turn.

Figure C.10: Model Validation - Gains from Marriage for Homogeneous Families by Province

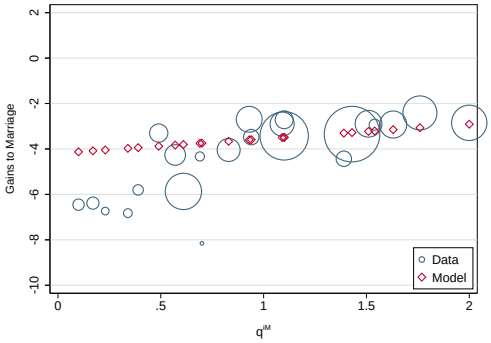
(a) Europe-EU15



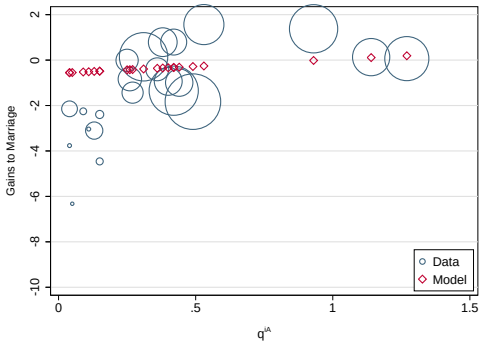
(b) Other Europe



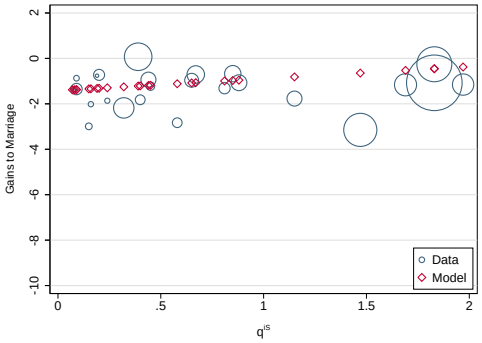
(c) North Africa-Middle East



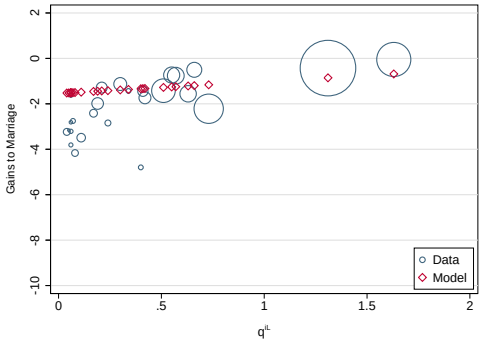
(d) Sub-Saharan Africa



(e) East Asia



(f) Latin America

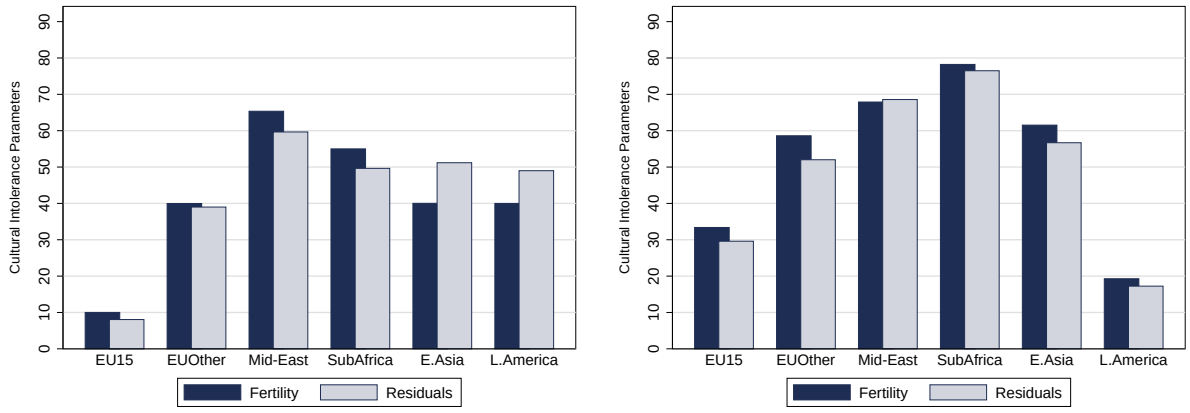


Notes: The figure shows out-of-sample predicted and implied gains to marriage for homogamous families of ethnic group minorities over the corresponding population share, q^i (in percentage), by province of residence (average over the time period). Empirical moments are weighted by the observed number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

Figure C.11: Cultural Intolerance Parameters - Robustness with Fertility Residuals

(a) Migrants towards Natives (Demand parameters)

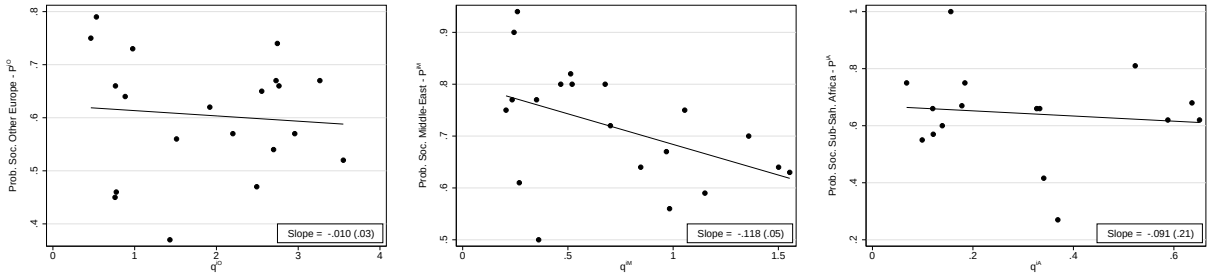
(b) Natives towards Migrants (Supply parameters)



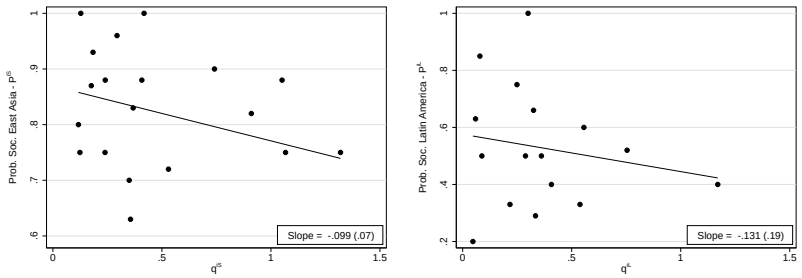
Notes: This figure reports parameter estimates of the cultural intolerance of immigrants versus natives ΔV_i^n (panel a) and natives versus immigrants ΔV_n^i (panel b) for all cultural-ethnic groups i. The blue bars report baseline estimates. The grey bars, instead, report estimates exploiting fertility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

Figure C.12: Minorities Socialization Probabilities and Horizontal Socialization

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

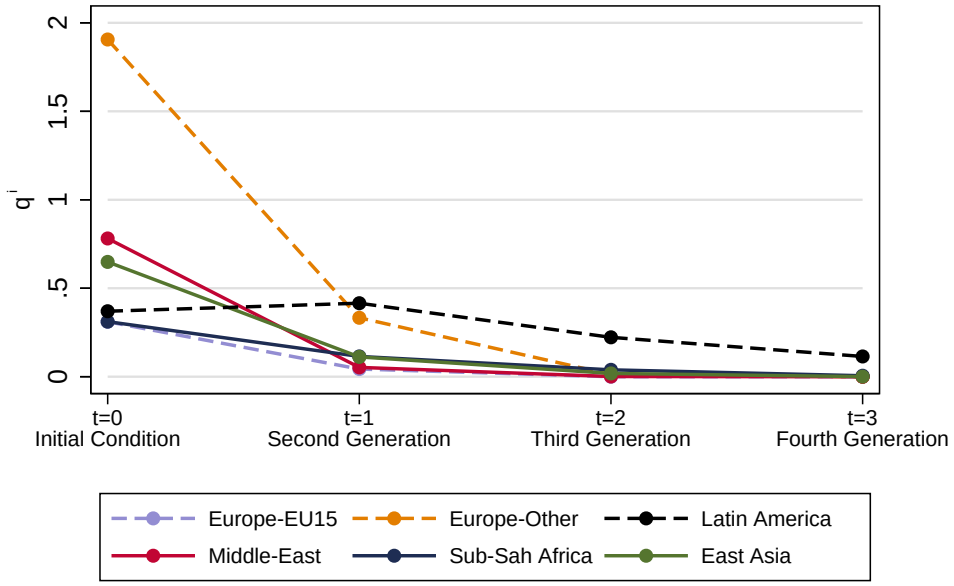


(d) East Asia (e) Latin America



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

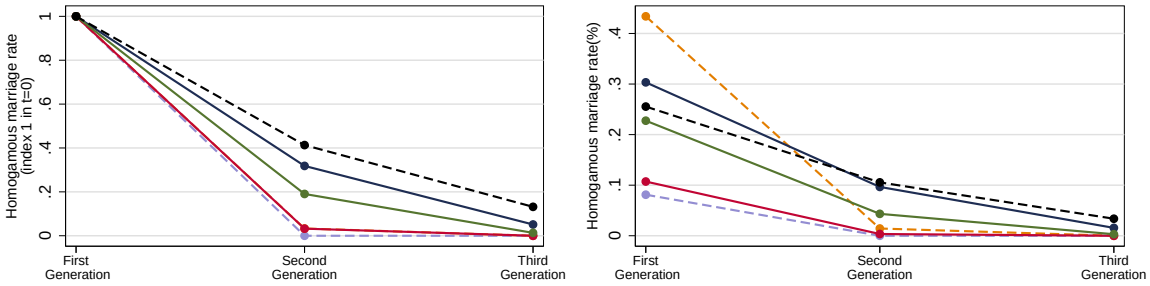
Figure C.13: Long-run Dynamics of Cultural Traits



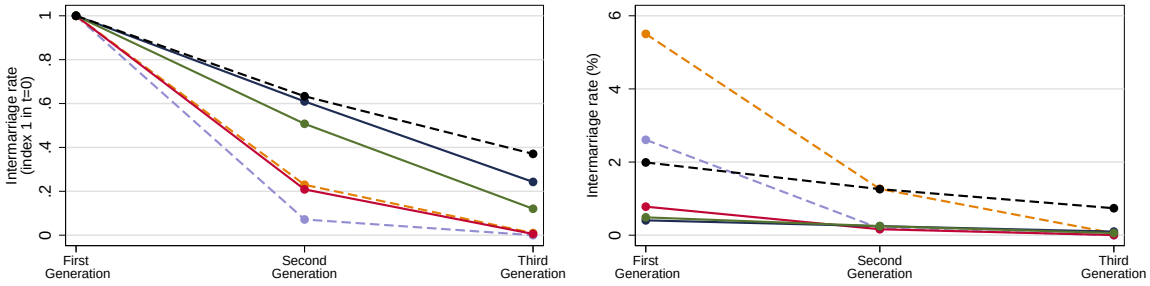
Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for all minorities, over successive generations.

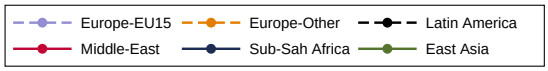
Figure C.14: Dynamics of Marital Matching

(a) Homogamous Marriages



(b) Heterogamous Marriages

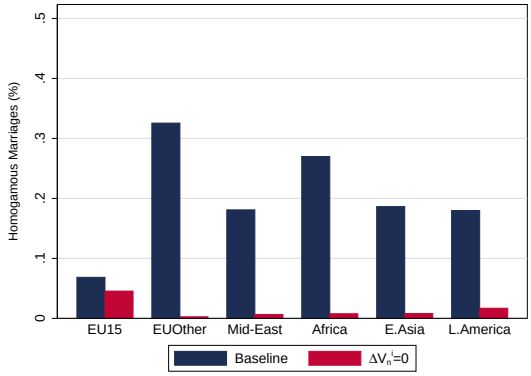




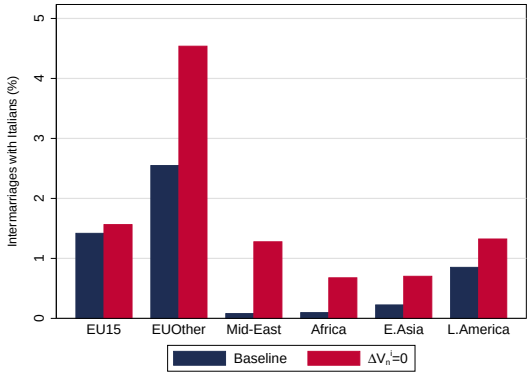
Notes: This figure shows the long-run dynamics of matching patterns for homogamous marriages (panel a) and heterogamous marriages with natives (panel b), over successive generations.

Figure C.15: Change in Matching Patterns with Italians Fully Tolerant, $\Delta V_n^i = 0$

(a) Homogamous marriages



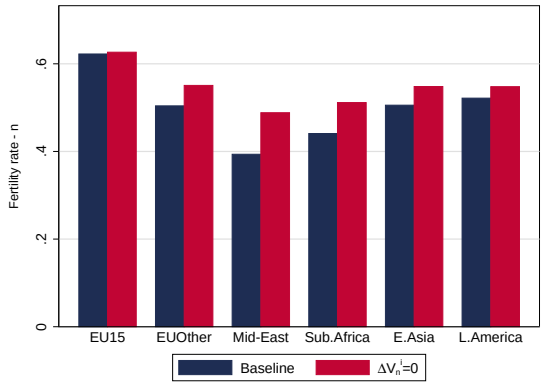
(b) Intermarriages with Italians



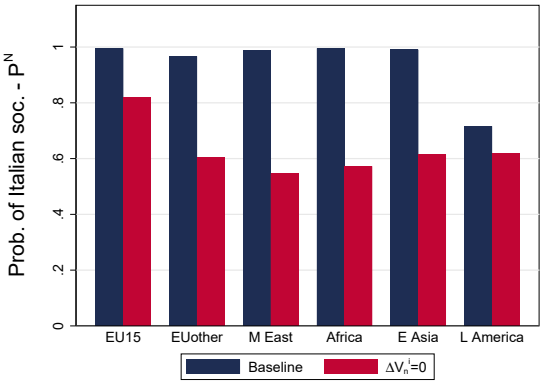
Notes: This figure shows the percentage change in homogamous (panel a) and heterogamous (panel b) marriages with full tolerance of natives towards minorities with respect to baseline.

Figure C.16: Change in Intra-household Patterns with Italians Fully Tolerant, $\Delta V_n^i = 0$

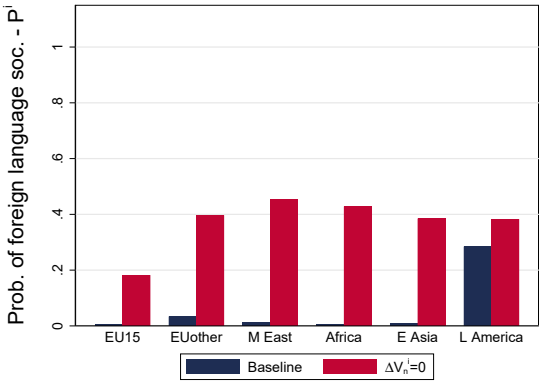
(a) Fertility



(b) Italian socialization

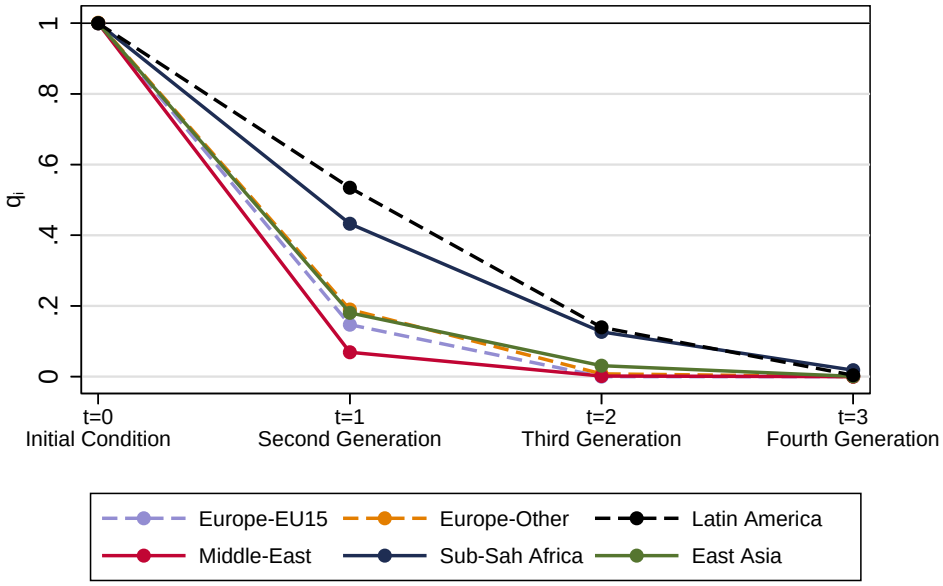


(c) Foreign language socialization



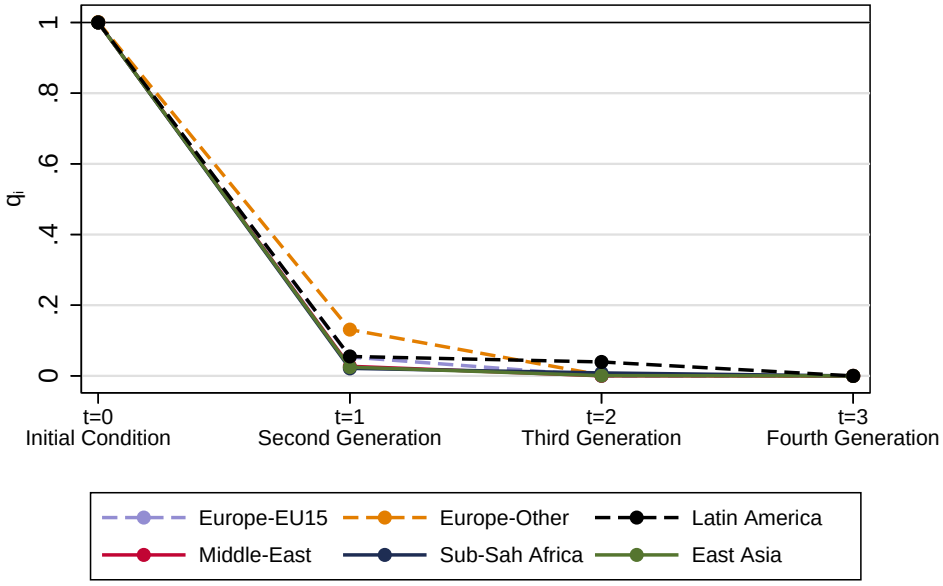
Notes: This figure shows the variation in fertility rate (panel a), Italian socialization probability (panel b) and foreign language socialization probability (panel c) in intermarriages with natives at the baseline and in case of complete tolerance of Italian majority towards minorities.

Figure C.17: Dynamics of Cultural Traits with Italians Fully Intolerant, $\Delta V_n^i = 100$



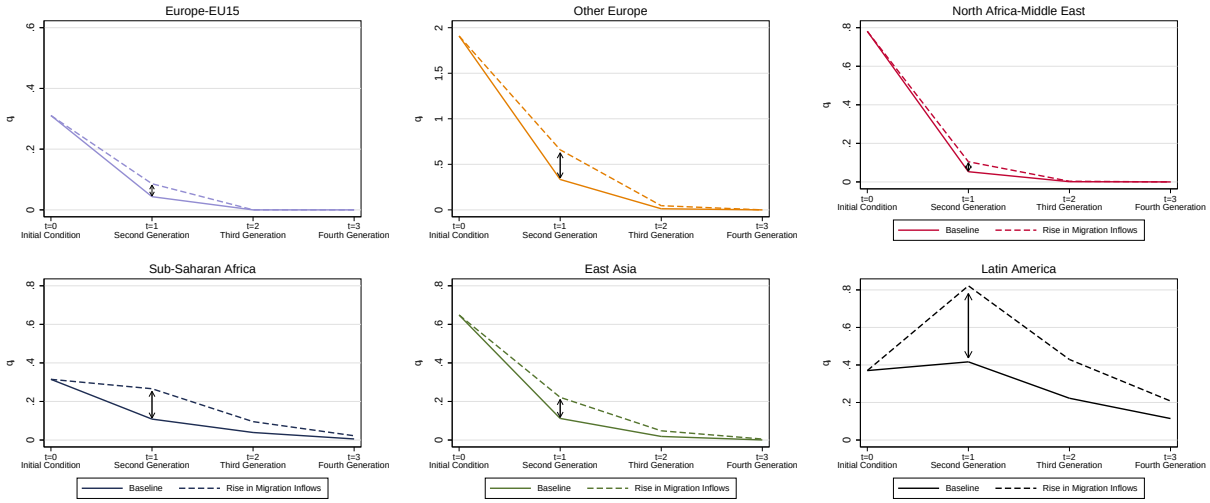
Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for all minorities i, over successive generations, assuming the case of complete intolerance of Italian majority towards minorities (q_t^i index to 1 in $t = 0$).

Figure C.18: Dynamics of Cultural Traits with Minorities Fully Tolerant, $\Delta V_i^n = 0$



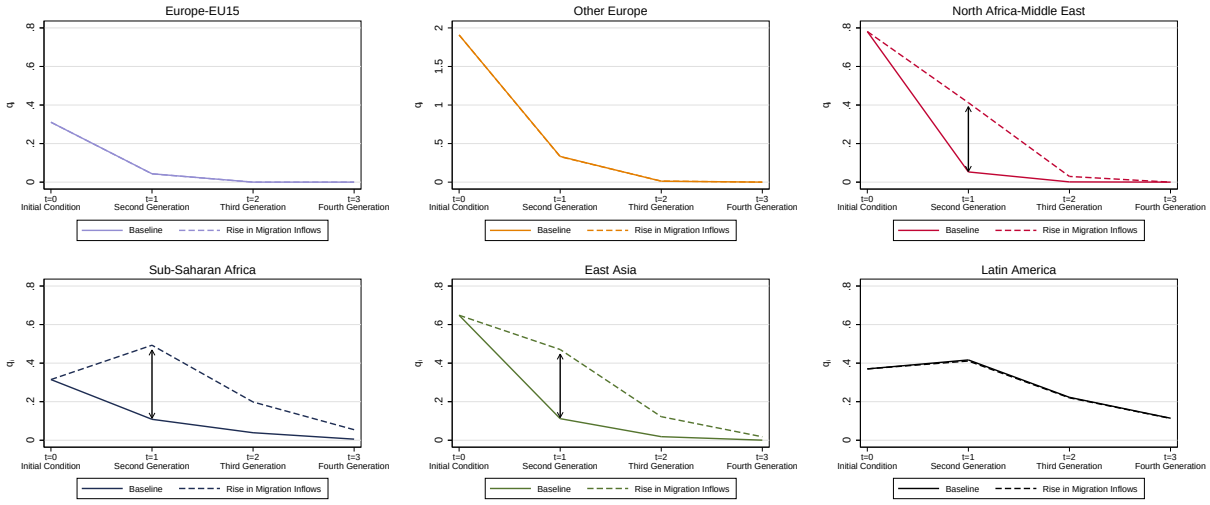
Notes: This figure shows the long-run dynamics of the distribution of cultural traits in the population for all minorities i, over successive generations assuming the case of complete tolerance of minorities towards Italian culture (q_t^i index to 1 in $t = 0$).

Figure C.19: Long-run Dynamics with Proportional Raise in Migration Inflows



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

Figure C.20: Long-run Dynamics with Raise in Specific Minorities Inflows



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

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

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

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

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

	Wife Ethnic Group:							
Husband Ethnic Group:	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah. Africa	East Asia	Latin America	Total
Italian	3,623,416	49,602	165,778	11,792	11,063	13,682	63,484	3,938,817
Europe-EU15	41,250	3,153	2,358	161	217	293	838	48,270
Other Europe	46,185	1,202	25,027	253	218	307	1,197	74,389
North Africa-Middle East	21,791	554	2,973	4,178	133	131	829	30,589
Sub-Saharan Africa	6,043	260	421	71	10,090	41	144	17,070
East Asia	2,420	127	348	51	44	9,865	129	12,984
Latin America	13,329	322	951	66	46	76	14,642	29,432
Total	3,754,434	55,220	197,856	16,572	21,811	24,395	81,263	4,151,551

Notes: This table reports the bivariate marriage distribution by cultural-ethnic group of spouses (absolute numbers).

Table C.3: Distribution of Singles by Cultural-Ethnic Group

	Panel a. Adult singles over 90th perc. of the Age at Marriage Distribution							
	Singles 2001				Singles 2011			
	Male	Share (%)	Female	Share (%)	Male	Share (%)	Female	Share (%)
Italian	2506182	18.57	5971291	34.71	3963745	26.13	7458287	40.03
Europe-EU15	19788	26.20	54228	32.53	43088	28.91	86383	42.11
Other Europe	25952	20.32	72587	36.63	86893	18.97	304691	39.04
Middle-East	16071	18.89	19940	35.37	34464	19.23	41189	36.63
Sub-Saharan Africa	5257	20.66	9641	38.54	15600	23.97	23905	42.44
East Asia	2886	12.53	9033	25.02	13949	15.24	36504	27.03
Latin America	11362	28.28	25875	35.86	31456	33.25	79113	43.25
	Panel b. Adult singles over 18 Years Old							
	Singles 2001				Singles 2011			
	Male	Share (%)	Female	Share (%)	Male	Share (%)	Female	Share (%)
Italian	7947039	36.87	9914990	42.42	8961649	41.29	11038623	47.18
Europe-EU15	84537	48.84	109512	40.05	86625	43.29	124133	46.33
Other Europe	124875	39.18	149279	36.87	312362	35.75	549604	40.31
North Africa-Middle East	61554	35.26	32328	28.59	106598	33.91	73237	30.20
Sub-Saharan Africa	24013	34.87	23711	41.83	58857	41.78	49560	44.94
East Asia	24819	31.54	23912	29.74	98240	34.24	84063	27.12
Latin America	33085	46.36	55992	41.64	84751	51.68	149838	49.23

Notes: This table reports the distribution of singles by gender and cultural-ethnic group, separately for 2001 and 2011. Panel a. reports the distribution of adult singles over the 90th percentile of the age at marriage distribution, and panel b. reports the distribution of adult singles over 18 years old. Shares are computed as the number of singles over the total number of individuals by gender and ethnic group, for 2001 and 2011 in turn.

Table C.4: 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 married and separated homogamous families, as well as married and separated heterogamous families. The separated category comprehends both separated and divorced unions.

Table C.5: Separation Rates by Ethnic Group of Spouses

	Separation Rates		
	Homogamous	Heterogamous	Heterogamous Italians excluded
Italian	0.064	0.075	-
Europe-EU15	0.024	0.048	0.058
Other Europe	0.030	0.071	0.057
North Africa-Middle East	0.045	0.116	0.070
Sub-Saharan Africa	0.026	0.092	0.066
East Asia	0.013	0.067	0.054
Latin America	0.050	0.092	0.076

Notes: This table reports the separation rates by ethnic group of spouses. Estimates are reported separately for homogamous, heterogamous, and heterogamous families excluding marriages with natives.

Table C.6: 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.7: Italian Language Socialization and Educational Outcomes - Robustness

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

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

Table C.8: Italian Language Socialization and Additional Measures of Integration

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

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

Table C.9: Structural Model Parameters -Robustness

Cultural Intolerance Parameters							
h:	Italian	Europe-EU15	Other Europe	Middle East	Sub-Sah Africa	East Asia	Latin America
ΔV_n^h, Italian		34.40	60.76	72.96	84.43	54.05	18.38
$\Delta V_{i_E}^h$, Europe-EU15	7.50		50.16	4.55	6.38	17.68	0.31
$\Delta V_{i_O}^h$, Other Europe	36.39	0.05		74.20	58.65	32.67	24.65
$\Delta V_{i_M}^h$, North Africa-Middle East	62.60	7.49	57.82		99.86	43.45	45.93
$\Delta V_{i_A}^h$, Sub-Saharan Africa	46.90	25.10	54.21	87.33		80.22	43.01
$\Delta V_{i_S}^h$, East Asia	38.48	0.30	75.82	50.70	44.63		43.42
$\Delta V_{i_L}^h$, Latin America	49.95	12.49	23.06	58.49	51.65	28.57	

Notes: This table shows structural parameter estimates, exploiting exogenous pre-determined variation in the distribution of population shares by ethnic group and region.