ALMA MATER STUDIORUM – UNIVERSITA' DI BOLOGNA

DOTTORATO DI RICERCA IN

ECONOMIA

CICLO XXIX

Settore Concorsuale di afferenza: 13 / A1

Settore Scientifico disciplinare: SECS - P01

Essays in the Economics of the Family and Cultural Transmission

Presentata da: Giulia Tura

Coordinatore Dottorato
Prof. Vincenzo Denicolò

Relatore
Prof. Giulio Zanella

Esame finale Anno 2017

Abstract

My research investigates the interplay between cultural transmission dynamics and household choices. I aim at identifying the role cultural-ethnic traits play in marriage choices and at uncovering the implications of marital sorting on consequent intra-household decisions. This thesis comprises four essays, which, from different perspectives, focus on interethnic marriages, within the Italian marriage market.

In the first chapter, *Marital Formation and Dissolution in Interethnic Marriages. Evidence from Italian Data*, I explore the role of ethnic endogamy on marital instability. Thanks to a unique quasi-longitudinal dataset, that I have constructed, matching marriage records with separation records from administrative Italian sources (1995-2012), I document the existence of a positive differential in marital dissolution of interethnic marriages as opposed to homogeneous ones. Results are heterogeneous by migrant ethnic group and I demonstrate that genetic and ethnolinguistic measures of cultural diversity between spouses are informative about the incidence of marital dissolution.

Starting from these findings, the second chapter *A Study of Marriage, Fertility and Divorce: Cultural and Ethnic Socialization of Migrants in Italy*, investigates a novel channel, which explains the differential in household stability and investments across family types, namely cultural socialization of children. I propose a marital matching model along cultural-ethnic lines, to study the process of family formation and intra-household decision making in a context where ethnic differences between spouses potentially matter both in terms of preferences and technologies for household production. I provide an explanation of several stylized facts in marriage markets, e.g., the strong positive assortative mating along cultural-ethnic identities and the relative lack of stability of inter-marriages, which centers around the role of cultural socialization of children. Taking advantage of rich administrative Italian data, I show that the observed intermarriage, fertility, separation and socialization rates are in line with theoretical predictions and they are consistent with strong preferences of parents toward cultural socialization of children to their own ethnic identity, proxied by language transmission.

In the third chapter, *The Price of Citizenship: The Effect of EU Enlargement on Marital Matching in Italy*, I propose and estimate a marital matching model along ethnic lines. I advance

the hypothesis that gains to intermarriage depend on both cultural preferences and citizenship or legal status acquisition motives. Taking advantage of the exogenous EU enlargements to East European countries in 2004 and 2007, I provide evidence that gains to intermarriage of East European migrants significantly decrease in response to the acquisition of a legal status . The decrease anticipates the introduction of the reform and remains persistent over time. Moreover, I find evidence in favour of cross-ethnic marital substitutability.

The final chapter, *Ethnic Judicial Bias: Discrimination or Integration? Evidence from Separation and Divorce Sentences*, aims to understand whether judicial decisions respond to the ethnic identity of spouses and what incentives those judgements are guided, by looking at family law proceedings. I analyse judicial decisions from the universe of separation and divorce sentences in Italy, from 2000 to 2012. Studying the legal custody assignment of children, I have documented that it is not the mother ethnic identity per se that affects her custody probability, but a significant differential is detected interacting mothers' ethnicities with the family type. New original data from transcripts of judges decisions, from the Family section of the Court of Milan, allows to improve the research, both methodologically and conceptually, in the ability to discriminate across different potential motivations that drive judges' sentences.

Contents

List of Figures

List of Tables

Introduction

A man is of all sorts of luggage the most difficult to be transported.
Adam Smith "Wealth of Nations"(1776)

Nowadays, economic literature has devoted large attention to the study of migration, both because of its increasing relevance in the last decades, but also because migration rises different issues concerning the integration of migrants into new communities, which is a complex process affecting the degree of inequality of our societies. How fast immigrants integrate into new communities from an economic perspective largely depends on how quickly they integrate culturally. In this respect, empirical research documents slow rates of convergence to natives culture and strong persistence in cultural traits from the part of minorities. Therefore, I believe, a deep investigation of the mechanisms that sustain cultural heterogeneity in the long run, is of major importance. Motivated by this reflection, my research studies the inter-generational process of cultural transmission, as centered on the role of the family. I interpret the family as the primary place where attitudes and beliefs are transmitted from parents to children. In particular, my thesis investigates the interplay between cultural transmission dynamics and household choices. I aim at identifying the role cultural traits play in marital formation choices and at uncovering the implications of marital sorting on consequent intra-household decisions.

Family economic literature has largely investigated the question of *who marries whom and why*, looking deeply into the assignment process of spouses, when competition for potential partners affects individual choices [Choo and Siow (2006b), Chiappori et al. (2012) and Dupuy and Galichon (2014)]. However, within these important contributions the cultural dimension has been greatly overlooked, despite the recognition of its relevance in explaining economic outcomes and of the strong persistence of cultural traits across generations [Bisin and Verdier (2000), Bisin et al. (2004) Fernández et al. (2004); Doepke and Zilibotti (2008); Fernández (2013)]. Building upon these premises, my research contributes to bridging the current gap between cultural and family economic literature, with the primary objective of understanding better the reasons why individuals select into the marriage market along cultural traits and values, and their implications. To this end, this thesis com-

prises four essays, which, from different perspectives, focus on interethnic marriages. I exploit variability in ethnic identity across migrants, within the Italian marriage market, in accordance with the continuing increase of the migratory phenomena.

In the first chapter, *Marital Formation and Dissolution in Interethnic Marriages. Evidence from Italian Data*, I empirically explore the role of ethnic endogamy on marital instability. I have constructed a unique quasi-longitudinal dataset, matching marriage records with separation records from administrative sources, covering the universe of marriages formed in Italy from 1995 to 2012. Compared to former sociological studies hampered by inherent data limitation and selection issues, I have exploited variability in marital duration and I have documented the existence of a positive differential in marital dissolution of interethnic marriages as opposed to homogeneous ones. Results are heterogeneous by migrants origin and I demonstrate that genetic and ethnolinguistic measures of cultural diversity are informative about the incidence of marital dissolution [Spolaore and Wacziarg (2009) and Melitz and Toubal (2014a)].

Starting from these findings, I enrich my analysis investigating marital matching, i.e. the selection of spouses into the marriage market along cultural lines. In the second chapter *A Study of Marriage, Fertility and Divorce: Cultural and Ethnic Socialization of Migrants in Italy*, joint with Alberto Bisin (New York University), I propose a structural model to study the process of family formation and intra-household decision making in a context where ethnic differences between spouses potentially matter both in terms of preferences and technologies for household production. I provide an explanation of several stylized facts we observe in marriage markets, e.g., the strong positive assortative mating along cultural-ethnic identities and the relative lack of stability of inter-marriages, which centers around the role of cultural socialization of children. I explicitly consider that parents care about the transmission of their own cultural traits to children and are endowed with technologies for cultural transmission. I propose to estimate model parameters on Italian data via a minimum distance procedure, taking advantage of rich administrative Italian data. Identification is achieved exploiting a multi-market framework, namely I exploit variability across marriage markets. In line with theoretical predictions, the observed intermarriage, fertility, separation and socialization rates are consistent with strong preferences of parents toward cultural socialization of children to their own ethnic identity, proxied by language transmission.

In the third chapter, *The Price of Citizenship: The Effect of EU Enlargement on Marital Matching in Italy*, I propose and estimate a transferable utility marital matching model along the cultural-ethnic identity of spouses. The analysis focuses on interethnic unions, between one Italian spouse and one non-Italian spouse, and exploits administrative individual level data on marriages formed between 1995 and 2012 in Italy. I argue that gains to interethnic marriage depend on both cultural preferences and citizenship acquisition motives. I

estimate the price of citizenship, taking advantage of the change in the EU citizenship regulation driven by the exogenous EU enlargement to East European countries in 2004 and 2007. Results show that legal status acquisition induced a huge decrease in the gains to intermarriage for migrants of new EU member countries, only after the EU enlargements. Heterogeneous effects across marriage markets are suggestive that intermarriages driven by legal status motives trade-off economic-labor opportunities. Finally, I provide evidence in favour of cross-ethnic marital substitutability in response to exogenous variation in legal status acquisition. A deep investigation of marital substitution patterns is a particularly important question. The project develops in this direction, proposing and estimating a formal multi-dimensional marital matching framework. I will increase the number of spouses relevant attributes for marital decisions, in terms of age and education.

The final chapter, *Racial Judicial Bias: Discrimination or Integration? Evidence from Separation and Divorce Sentences*, aims to understand whether judicial decisions respond to the ethnic identity of spouses and what incentives those judgements are guided, by looking at family law proceedings. I analyse judicial decisions from the universe of separation and divorce sentences in Italy, between 2000 and 2012. Studying the legal custody assignment of children, I have documented that it is not the mother ethnic identity per se that affects her custody probability, but a significant differential is detected interacting mothers' ethnicities with the family type. For instance, compared to native mothers in homogeneous marriages, foreign mothers in mixed families are significantly less likely to obtain their children custody, while foreign mothers in homogeneous families are favored. Estimates are heterogeneous across ethnic groups. The project takes advantage of the random assignment of cases to judges to rule out potential unobserved case heterogeneity and it will provide between-judges evidence of ethnic judicial disparities. New original data from transcripts of judges decisions, which I am currently collecting from the Family section of the Court of Milan, allows to improve the research, both methodologically and conceptually, in the ability to discriminate across different potential motivations that drive judges' sentences.

Chapter 1

Marital Formation and Dissolution of Interethnic Marriages. Evidence from Italian Data [1]

Abstract

This research uses unique administrative Italian data (1995-2012) to investigate the role of endogamy in the processes of marriage formation and dissolution. It focuses on interethnic marriages and follows the evolution of the migratory inflows that have characterized the Italian territory over the last decades. The analysis provides updated evidence of the existence of a positive differential in marital dissolution in interethnic marriages as averse to homogeneous ones. The results also suggest the presence of a foreign spouse gender effect, i.e. the separation risk is higher in mixed families with a non-Italian husband than those with a non-Italian wife. Asymmetries observed among native and foreign mothers in the probability of children custody assignment might account for differences in mothers' separation choices. Finally, the results are heterogeneous by origin of migrants. Genetic and ethnolinguistic measures of cultural diversity are informative about the incidence of marital dissolution, i.e. a standard deviation increase in the ethnic diversity of spouses induces a 5-8% increase in marital dissolution risk.

Keywords: Ethnic intermarriage, Marital Formation and Dissolution, Cultural distance.
JEL Classification: C41, J12, J15, J16.

[1] This chapter is based on ideas discussed in 2014 with Laura Bottazzi and Paolo Manasse, who were then working at a similar project, I am grateful for the exchange. I wish to thank Marina Briolini and Sara Piombo at the ADELE Laboratory in Bologna for their support in providing me useful documentations and suggestions.

1.1 Introduction

There is a growing body of literature in family economics that studies the aggregate patterns of family formation. Family economics literature began with Gary Becker's seminal contributions [1973, 1974][2] where he transferred the concept of supply and demand analysis to the marriage market. He examined the household formation and dissolution choices, assuming that the marriage is a voluntary choice of utility maximizing agents, and assuming also that the marriage market is in equilibrium. Other literature that followed Becker's work, further investigated who marries whom and why [Choo and Siow (2006b)] and looked more deeply into the assignment processes of spouses, when competition for potential partners in the marriage market affects individual choices. For instance, the decision to enter into a specific union depends on the value of the match, but also on the whole range of potential matches in the market [3]. The matching literature was initially concerned with investigating how changes in preferences or characteristics in the population vector affected the equilibrium assignment in the market, along the various sorting dimensions. Some studies put a particular focus on the sorting mechanisms of: age [Choo and Siow (2006b) and Choo (2015)], income and body mass index [Chiappori et al. (2012)]; education and socio-economic status [Chiappori, Oreffice, and Quintana-Domeque (Chiappori et al.) and Chiappori, Salanié, and Weiss (Chiappori et al.)] and personality traits [Dupuy and Galichon (2014)]. However, differences in cultural values of spouses have been far less investigated. More recently, the research has explored the relationship between marital assignment and the following household dimensions. There were various reasons for this increased focus. First, individuals select into a marriage by internalizing the future value of their marriage, so the observation of assortative mating gives insights into intrahousehold relationships, technologies and allocation decision (sharing-rules). Secondly, by studying the evolution in matching patterns, over time, an understanding can be gained about the changes that are brought about, in the household, by economic, cultural and legal changes. Consider, for example, how the mitigation of the stigma of divorce, and the reduction in the legal costs required to obtain it, have contributed to a modification in the marital dissolution rate. Moreover, marital sorting was investigated as a potential source reinforcing between households inequality and strengthening the long-term social boundaries, that can be caused by intergenerational transmission [Fernández et al. (2005) and Greenwood et al. (2014)]. Finally, marital matching affects marital stability because a lower quality of sorting affects the marital outcomes variability [Becker et al. (1977)].

[2] Becker (1973); Becker (1974). See also Becker et al. (1977) and Becker (1991).

[3] Relevant related references: Fernández et al. (2005); Choo (2015); Siow (2015); Bruze et al. (2015). For a survey on the marital matching literature see also Browning et al. (2014) Chapter 7 and Chiappori and Salanie (2016).

The present research specifically explores the latter dimension: marital instability. I have studied the role of endogamy (i.e. marrying within the same ethnic group) in characterizing the process of marriage formation and dissolution. I focused my research on one matching dimension, notably the cultural-ethnic identity of the spouses. In this paper, I argue that the spouse's ethnic origins are complement inputs within the household production function. As a consequence, the higher the cultural distance between the spouses, the lower the expected quality of the marital union. In expectation, higher uncertainty about the marital outcomes reduces the incentives for making specific investments in the family, which also increase the likelihood of the marital dissolution. I present evidence that shows that there is a higher risk of separation among interethnic families than in co-ethnic ones. Moreover, my findings suggest that there is a positive relationship between the cultural distance between spouses and the risk of marital dissolution. This hypothesis rests on two facts. First, the literature in different fields explains that spouses' cultural roots show complementarities in the household production function. In particular, the sociological literature stresses the role played by common language in facilitating communication and understanding and therefore investment decisions [Stevens and Schoen (1988) and Kalmijn et al. (2005)]. Moreover, the literature concerning cultural transmission underlines the point that homogamous unions benefit from coordinated incentives, especially, in the cultural socialization of the children with respect to heterogamous families, hence they have access to more efficient socialization technologies [Bisin and Verdier (2000)]. Secondly, as documented in Becker's original writings, complementarities in production inputs lead to positive sorting in the marriage market, i.e. positive assortative mating (PAM). Marital dissolution choices are thus rationalized, introducing uncertainty in marital outcomes. Higher uncertainty about the outcome of the marriage, caused by a lower quality in the initial marital matching, will increase the likelihood of the marriage's being dissolved.

In the following article, I provide evidence to support the *heterogamous hypothesis*, building my findings on an exploration of ethic intermarriages celebrated in Italy. I exploit within country variability in the cultural identity of migrants, according to the evolution of the migratory inflows that have characterized the Italian territory over the last decades. Indeed, Italy experiences different waves of immigration with heterogeneity both in time and ethnic composition. At the beginning of 2014, the number of foreign citizens registered in Italy was more than 4.9 million and they accounted for 8.1% of the total resident population. This is in comparison to 2013, when the same percentage was around 7.4% and as little as ten years earlier, in 2003, it was 2.6%. I provide evidence that these migratory inflows have affected the Italian marriage market; the percentage of interethnic marriages nearly tripled during the period of my investigation. The richness of the available data allowed me to go beyond the racial differences that have been previously investigated and where patterns of marital instability may also have suffered from the socioeconomic marginalization concerns.

My analysis exploits unique Italian administrative data that covers the universe of marriages celebrated in Italy, from 1995 to 2012 and which corresponds to the period of systematic increase in immigration. Having access to the actual data on the timing of marital dissolution [4], I estimated the role of ethnic assortative mating on marital instability using a Cox PH duration model. This empirical analysis provides evidence of a positive differential in the marital instability of heterogamous marriages as compared to culturally homogeneous unions. I show that the differential remains large and statistically significant, after matching marriages with respect to all the relevant assortative mating dimensions that might correlate with cultural sorting choices, such as age, past marital history, education and labor choices [Becker et al. (1977); Kalmijn et al. (2005); Choo and Siow (2006b), Frimmel et al. (2013); Choo (2015) and Bruze et al. (2015)]. Results also suggest the presence of a foreign spouse gender effect, i.e. the risk of separation is 6.4% higher in families that have a non-Italian husband than in heterogamous families where the wife is non-Italian. This differential could be explained by the estimated difference between native and foreign mothers in the probability of children custody assignment. Moreover, when one allows for heterogeneous effects in the generational order of migrants, the evidence highlights a tendency towards migrants assimilation into the host country, in terms of family dissolution decisions. Finally, I show that there is a positive correlation between instability and cultural distance by the macro area of origin. Delving further into this dimension, I investigate the effect of ethnic-cultural differences, between spouses, on the risk marital dissolution. I exploit genetic and ethnolinguistic distance measures as proxies for the spouses' cultural diversity [Spolaore and Wacziarg (2009), Guiso et al. (2009); Michalopoulos (2012) and Melitz and Toubal (2014b)]. Evidence shows that genetic and ethnolinguistic measures of cultural distance are informative about the marital dissolution incidence, i.e. a standard deviation increase in the cultural diversity of spouses induces a 5-8% increase in the risk of marital dissolution.

The research aims to contribute to the family economics literature, in providing evidence of decreasing assortativeness along cultural-ethnic traits. It is beyond the scope of this project to present reasons for this increase in heterogamy, which would require more in-depth research. For instance, changes in preferences for heterogamy might be the driving factor behind the observed increase in interethnic marriages, but this could just as equally come from changes in population vectors or, more generally, from the process of the migrant' integration, lowering the perceptions of cultural distances. Moreover, I aim to contribute to the literature on culture. There is a huge and expanding body of literature that exploits within country cultural heterogeneity by ethnic and religious traits [Fernández and

[4] Interestingly data allows to focus both on Separations and Divorces, which correspond to two different legal processes within the Italian institutional framework. Empirical analysis is primarily based on separations data.

Fogli (2009); Bisin and Verdier (2011)], to investigate the role of culture in shaping economic outcomes. In general, all those contributions uncovered a strong resilience of cultural traits. How those traits affect marital decisions and interactions inside the household is a very important question in ongoing research. Finally, interethnic marriages have been studied, extensively, within migration literature as a way of evaluating the process of cultural assimilation and the integration of migrants into the host country [Meng and Gregory (2005); Constant and Zimmermann (2008); Constant et al. (2009) and Algan et al. (2012)].

In the following section, I provide an account of the existing literature concerning how families are formed and their dissolution patterns. In Section 1.3 the data is introduced and the empirical strategy is explained. Section 3.6 covers the empirical findings about family formation and the results of the duration analysis are presented in Section 1.5. Finally, Section 3.7 concludes.

1.2 Literature Review

In the last decades economic research have studied the evolution of the family as an institution. They have mainly investigated the reasons for, and the consequences of, the changes that occur over time and that affect the process of family formation and dissolution [Browning et al. (2014)]. Among all the changes experienced by the family, particular attention has been devoted to the increasing likelihood of divorce; both because of its consequences and its importance. On one hand, empirical research is involved in the evaluation of the effects of divorce on different outcomes, explotinig exogenous changes in divorce or custody law over time and across states [Amato (1993); Amato (1996); Friedberg (1998); Amato (2000); Stevenson and Wolfers (2006); Stevenson and Wolfers (2006); Fernández and Wong (2014); Del Boca et al. (2014) and Voena (2015)]. Studies have shed light on the effects of the increasing marital dissolution probability on intra household bargaining, the women labor force participation, the well-being of children, the decisions concerning investments in children's education and on the intergenerational effects of offsprings' marital decision. On the other hand, a parallel body of literature has investigated the causes of divorce, even if causal identification is difficult to reach in this context, because of self-selection, both in marital formation and dissolution. A better understanding of the drivers that induce families formation and that cause their dissolution is thus important in order to understand how they feed back into economic decisions.

The following section concentrates, in particular, on the implications endogamy has on decisions concerning marital formation and dissolution. Nowadays, a large literature in different field investigates the subject. Sociological literature describes endogamy as the result of three different potential sources. First, it might come from an individual's preference for a mate with similar characteristics to themselves. Secondly, it might be driven by the

interference of third parties; mostly the social or parental influences might govern the processes of marriage formation. Finally, geographical constraints might induce homogamy, by limiting contact with potential culturally different spouses, i.e. residential, educational or labor segregation[Rosenfeld (2008)]. Economic literature mainly follows the first direction, modelling individual's preferences for partners characteristics [5]. Beginning with the assumption that there is complementary of inputs in the household's production function, Becker (1973)'s contribution predicts that positive assortative mating is optimal, i.e. individuals will optimally select partners with similar characteristics, in terms of intelligence, race, religion, education and wage rate, in order to profit from complementarities in the household's production capabilities [6]. Becker et al. (1977) elaborate this further and propose a theoretical analysis of marital dissolution, introducing uncertainty about outcomes of marital decisions, i.e. a deviation between the expected and realized utilities. Uncertainty makes dissolution choices stochastic with their probability depending on two components: the expected gain from marriage and the distribution of unexpected outcomes. Thus, the stability of marriage is a function of the expected quality of the sorting process. Consequently, the likelihood of divorce becomes smaller as the expected gains from marriage become larger, which are positively related to the quality of the marital matching, and the lower it is the variance of the distribution of unexpected gains from marriage.

From an empirical perspective, Becker et al. (1977) investigated the probability of marital dissolution, using cross-sectional data from the Survey of Economic Opportunity (SEO US Bureau of the Census, 1967). Their results confirmed theoretical model predictions. First, where mixed-religion couples were concerned, the research showed that the lower the sorting quality, the higher the dissolution probability and the lower the probability of remarriage if divorced. Secondly, they found that race differences reduced marital specific investments (i.e. a lower rate of childbirth), which in turn increased the probability of dissolution. The direction of the effect might be also reversed, i.e. a higher probability of dissolution, provoked by religious and race differences, decreased the investment in marital specific capital [7].

[5] Concerning parental influence, see Huang et al. (2012).

[6] Becker (1973) postulates the existence of a household production function, hence individuals maximize the share of consumption derived from household-produced commodities. Household formation choices depend upon the comparison of the gains from marriage with respect to remain single, where marriage gains are function of income, human capital and relative difference in wage rate of spouses. In particular, the model derives a positive correlation between spouses dimensions that enter as complementary inputs in the household production function, while a negative correlation is implied between inputs that are substitutes in the household production. Hence similarity between spouses is optimal as far as it enhances complementarity in household production and consumption. Recent contributions propose models for the empirical validation of the positive assortative mating hypothesis estimating explicitly a marriage matching function, both in static and dynamic stochastic environments [Choo and Siow (2006b); Siow (2015); Bruze et al. (2015) and Choo (2015)].

[7] On the same line Weiss and Willis (1997) study how unexpected shock in observable characteristics of either spouse affect marital dissolution hazard conditioning on observable characteristics of the partners. The main contribution is that they disentangle the divorce probability related to bad initial sorting with respect to

Becker's studies in family economics inspired the literature that follows. In particular, all the empirical investigations universally confirm that heterogamous marriages experience lower stability and a higher risk of divorce with respect to homogeneous marriages, along different assortative mating dimensions such as age and education [Weiss and Willis (1997); Greenwood et al. (2014); Bruze et al. (2015) and Choo (2015)] and religious affiliation and ethnic identity [Jones (1996), Kalmijn et al. (2005), Bratter and King (2008) and Zhang and Van Hook (2009)]. Most of the existing literature has explored the US environment and used differences across races, beginning with white-black marriages [Monahan (1970)] [8]. Bratter and King (2008) focused on the stability of interracial marriages in the US and documented that the percentage of couples divorcing after ten years of marriage was higher among interracial couples than same race couples, especially among the more recent marriage groups. In an interesting study, Phillips and Sweeney (2006) investigated how different risk factors might affect marital dissolution, for Mexican women. They decomposed the differences in marital dissolution probabilities between US-native and migrants Mexican women, into a component that accounted for the differences in means of the risk factors (US-native Mexican women have more education, are more likely to work, to have co-habited or to have had a child before marriage, than migrant Mexican women) and other (unexplained) factors. They found that the probability of dissolution, among US-born Mexican women, diminished by 25% if they had the same compositional characteristics as Mexican-born women. Compared to past studies, Zhang and Van Hook (2009) provided evidence about dissolution patterns in mixed marriages, using more recent US data, from 1990 to 2001. They found that interracial marriages are less stable than endogamous marriages, with a differential of about 15% in the risk of divorce. However, by using race and ethnicity variability for the most represented migrant groups, their results revealed asymmetries in marital instability. As an example, mixed marriages involving Blacks were the least stable followed by Hispanics, whereas mixed marriages involving Asians were even more stable than endogamous White marriages. They interpreted the evidence as a by-product of the specific minorities' difficulties in integration. The social isolation of Black-White unions, especially from the White community, threatened the stability of these marriages [Porterfield (1982) and Yancey and Yancy (2007)]. Jones (1996) enlarged the analysis to Australian and Hawaiian divorce data, in order to study the convergence behavior of ethnically mixed couples in terms of their divorce decisions. In particular the author showed that the risk of divorce, among mixed groups, falls in between the divorce patterns of the involved groups.

unexpected shocks, exploiting the National Longitudinal Study of th High School Class of 1972. Controlling for matching quality in observable characteristics, they confirm that similarity in ethnicity and religious belonging between spouses are shown to negatively influence the divorce hazard, so it increases marital stability.

[8] Refer to Fryer (2007) for deep description and investigation of the trends in interracial marriage in US from 1880 to 2000.

Despite the fact that the question has received a great deal of consideration, few contributions have investigated racial and ethnic differences on family patterns in European countries. As an exception; Kalmijn et al. (2005) exploited marriage data (1974-1984) from the Netherlands and demonstrated that a modest relationship existed between religious heterogamy and divorce, while a strong relationship was evidenced between national heterogamy and divorce. In particular, the marriages of Dutch spouses to foreigners had an average unconditional divorce risk that was twice as high as the maximum level of divorce in the two groups of reference. They distinguished between Western European, Southern European, Turkish and Moroccan migrants and showed that the heterogamous effect was stronger, the higher the cultural dissimilarity of the migrants from the natives. This suggested that newer ethnic boundaries are becoming increasingly more important in society than the older religious ones. Unfortunately the limitations of the data do not allow the quality of the matching, in the other relevant observable dimensions, to be controlled. Hence, the results have to be interpreted as an upper bound of a genuine relation, assuming a positive correlation between unobserved matching dimensions and ethnic ones. Finally, it is important to stress that very little is known about interethnic marriage instability in contemporary European society. Taking the radical and continuous changes in the family, in recent times, into account, this research aims to fill this gap by providing updated and comprehensive empirical conclusions. This descriptive paper aims to highlight a set of stylized facts that should rank highly in the research agenda on family and cultural economics.

1.3 Data and Estimation Strategy

1.3.1 Data

The empirical analysis is based on unique and high-quality administrative data on Marriages, Separations and Divorces, collected from two different sources and provided by ISTAT (National Statistical Institute). Specifically, Marriage datasets come from the municipal *vital statistics' registries* and cover the universe of marriages celebrated in Italy from 1995 to 2012. In addition, Separation and Divorce datasets come from the *registries of civil court chancelleries* and cover the universe of separations and divorces registered in Italy in the same period (1995-2012) [9].

[9] Marriages administrative micro-data comes from questionnaires completed by the registrar of the municipality where the marriage was celebrated. The micro-data about separations and divorces comes from questionnaires completed directly by the courts' chancelleries. The fact that information comes from objective sources (i.e. information does not come from individual self-reporting or retrospective acknowledge) increases the reliability of the analysis. For the period under investigation, registries of civil court chancelleries constitute a unique source for separations and divorces data, while starting from December 2014 (in application of Law 162/2014) consensual separation and divorce proceedings can be submitted to the civic registrar. This excludes potential issues in the selection of available data.

The *Marriage* datasets include information on (i) the personal characteristics of the spouses at marriage, such as age, educational attainment, professional status and occupation, the Italian municipality of residence and of birth, whether the spouses were born in Italy or not and their relative citizenship information and (ii) marriage information such as the date and place of the marriage and whether spouses choose a joint prenuptial contract or not. The datasets on *Separations* and *Divorces* contain details such as (i) the same personal characteristics of spouses as previously mentioned, (ii) retrospective information on marriage and family characteristics such as the date of the marriage, the place of residence, the number of children born in the marriage and the demographic characteristics of minor children present in the household, (iii) information on the separation and divorce proceedings in terms of the type of procedure (consensual or judicial), the type of conclusion of the proceeding (conciliation, approval, cancellation, court acceptance, judge remittance or change of proceeding) and (iv) the post-dissolution arrangements such as alimony obligations, recipient subject and annual amount of alimony, the assignment of child custody and the disposition of the family home [10]. Interestingly data allow to focus on both Separations and Divorces, which correspond to two different legal processes.

The empirical estimation is based on a unique dataset that combines information on marriages and dissolutions from the above mentioned sources and that covers the universe of marriages celebrated in Italy from 1995 to 2012. The final sample consists of 4,462,229 marriages, with 7% of them having been dissolved during the sample period. Time invariant dimensions were used to link marriages from different sources, i.e. the exact date of the marriage and the exact date and province of birth of both spouses. The combination of these characteristics meant that only 0.5% of the marriages were dropped from the sample, suggesting that marriages can be uniquely identified through this set of characteristics [11]. The design allows to access the risk of marital dissolution in the first years of marriage. In order to account for out-migration choices of families, the sample is restricted to marriages where at least one spouse was resident in Italy at the moment of the marriage. Moreover to control for mortality issue, estimates by age category are presented as robustness. Results

[10] Proceedings classified to end with conciliation, cancellation, or change of rite are registered, but no information is available for them. They have been dropped from the final sample because not representative of effective marital dissolution choices. Separation records that end up in conciliation are 2,149, those cancelled are 18,084 and those that changed rite are 1,772; hence they account of the 1.59% of the total number of separations. While for divorces, evidence reports that 1,067 proceedings end up in conciliation, 2,168 cancelled and 203 changed rite; for on overall 0.43% of the total of divorces.

[11] The merge procedure might suffer from coding errors. Moreover separations related to marriages celebrated abroad cannot be traced in vital records registries and hence excluded from the sample. Among the not-merged separations, the number of heterogamous families is overrepresented, as expected, because those couples are more likely to have celebrated their marriage abroad. Despite the possibility of deriving their marriage duration from separation information data, their inclusion in the sample might bias estimates because of truncation spell-selection. Right truncation leads to an over-representation of short spells relative to long spells, i.e. for all people beginning a marriage at a particular date, those who are more likely to survive are less likely to be found in the outflow at a particular date. The selection is symmetrical in case of left-truncation.

remain consistent.

The study of the Italian context is particularly interesting, for two main reasons. First, Italy legalized divorce only quite recently compared with other Western countries. Thus the evaluation of the marriage dissolution pattern over the last two decades is particularly relevant [12]. Secondly, the focus on interethnic marriages is in line with the evolution of the migratory phenomena that have characterized the Italian context starting at the end of the eighties when Italy began to undergo different waves of heterogeneous immigration, both in time and ethnic composition [13]. The inward flows of migration affected the Italian marriage market and increased the incidence of interethnic marriages. For instance, in 1995, 95.75% of all marriages celebrated in Italy were amongst native Italians; however the percentage decreased to 86.28% over the following ten years. Symmetrically, the percentage of mixed marriages (between a native Italian and a foreigner) celebrated in 2012, equaled 10.45% of the total number of marriages, while 5.06% were marriages between foreign spouses.

For the empirical purpose, differences in cultural traits are evaluated with respect to the ethnic identity, proxied by the country of origin. It is worth stressing that ethnicity is universally interpreted as an inherited trait, i.e. its evolution is defined by intergenerational transmission as averse to acquired traits, such as schooling. Both dimensions influence marriage patterns, but while the acquired traits are affected by the perspective of marriage, the inherited traits are independent of these feedback effects.

The fact that cultural values are relevant is supported by the observation of a strong resilience in ethnic traits. One clear manifestation of this phenomenon is cultural heterogeneity and the geographical fractionalization by ethnic and religious traits over time. For instance, Fernández and Fogli (2009) provided evidence in favor of a strong resilience in ethnic belonging among second generation migrants in the US, in terms of female labour participation and childbirth choices. Bisin et al. (2016) document that the strength of ethnic identity is higher in mixed rather than segregated neighborhoods, which highlighted the

[12] Divorce was legalized in Italy by Law 898/1970. After 1970, only two major reforms intervene to modify it. The first reform (Law 436/1978) improves the economic protection of the weaker partner guaranteeing heath insurance assistance. The second reform (Law 74/1987) lowers the separation period requested before a divorce application (from 5 to 3 years). Changes in the legal structure of separation and divorce have been shown to increase the marital dissolution incidence [Friedberg (1998)]. During the period under investigation, only one relevant divorce law took place in 2006 (Law 54/2006), which changed the standard arrangement from sole custody to joint custody of minor children. A detailed description of the law provisions is presented by De Blasio and Vuri (2013).

[13] The percentage of foreigners over the total resident population in Italy, 7.4%, is slightly below the EU average of 8.0% (2013), not too far from UK (7.7%) and Germany (9.4%) and higher with respect to France (6.2%) where citizenship acquisition rules are different and the naturalization process is more pronounced. Figures A.5 presents the evolution of the migration inflows by country of origin for the most represented ethnic groups (2002-2012). It underlines both the evolution of the immigration phenomena and the composition changes over time Data on total migration inflows are provided by Istat Survey: Movement and Annual Evaluation of the Foreign Resident Population by Origin Classification.

fact that the cultural distinction mechanism prevails over conformist preferences [14].

The ethnic identity of spouses is measured by their place of birth, therefore it is possible to identify native and migrant spouses and for the latter group, to distinguish first and second (or further) generation migrants. Robustness, using citizenship information, have been performed to strength the comparison with previous researches. Results are robust to different ethnic classifications. Nevertheless citizenship classification might overestimate the role of ethnic assortative mating on marital dissolution because of naturalization policies [Kalmijn et al. (2005) and Frimmel et al. (2013)]. Defined the ethnic identity of individuals, families are divided in three categories: *Homogamous* families where both spouses are Italian natives; *Heterogamous* families where one spouse is a native Italian and the other one is a migrant (indifferently for first and second generations) and *Migrants* families when both spouses are migrants. Heterogamous families are further divided, based on the gender of the migrant spouse, into *Heterogamous Wife* where the wife is born outside of Italy, and conversely *Heterogamous Husband* for families where the husband is the foreign-born spouse. Finally, because of the regulations on the acquisition of Italian citizenship, it is possible to classify migrants in respect of their generational order, into first and second generation migrants, using the available citizenship acquisition information (Italian born, naturalized Italian and non-Italian). This generation classification allows insights to be drawn concerning the assimilation processes of migrants into Italy. The classification leads to five different categories of family type: *Heterogamous First, Heterogamous Second, Migrants First, Migrants Second and Migrants Mixed*.

Empirical evidence suggests that the percentage of homogamous marriages, out of the total number of marriages in a given year, lowered over time, while the equivalent percentage of heterogamous and migrants' marriages increased over time[15]. More interestingly, by looking only at marriages celebrated from 1995 to 2012, a comparison of the marriage and separation rates shows that the incidence of dissolution is higher in heterogamous families than in homogamous ones.

Finally, information about the country of origin of the migrants is used to derive the

[14] The interest for the ethnic dimension with respect to other cultural dimensions (e.g. religious affiliation or religious intensity) is motivated by different reasons. First of all, the secularization process and the increase in migration inflows in Italy, make religious boundaries less salient with respect to ethnic ones [Dalla Zuanna (2008)], as observed also in other Western countries [Kalmijn et al. (2005)]. Secondly, religiosity is generally measured through subjective reporting of denomination suffering from potential measurement errors: subjective reporting might reflect not only actual religiosity but also social acceptance; and retrospective information is influenced by actual behaviours and life-course choices. Finally, interethnic marriages are extensively studied within migration literature as a dimension to analyse the process of cultural assimilation and integration of migrants in the host country [Meng and Gregory (2005); Constant and Zimmermann (2008); Constant et al. (2009) and Algan et al. (2012)]

[15] Descriptives on the absolute number of marriages, separations and divorces by family type are reported in Table **??** for the entire investigation period (1995-2012) and specifically every five years (1995, 2000, 2005, 2010) to draw insight of the pattern of the phenomena over time.

cultural distance between the spouses, for the years from 2005-2012. The distance in the cultural origins of the spouses is measured by the following: (i) the genetic distance and (ii) the ethnolinguistic distance.

Genetic distance measures the degree of genealogical relatedness between two populations and it is associated to the time elapsed since two populations' last common ancestors [Spolaore and Wacziarg (2009)]. The metrics for genetic distance is based on the genetic tree classification, as described in Cavalli-Sforza and Piazza (1994). The first genetic distance measure used (*Genetic distance I*) is defined based on the *coancestry coefficients*: the heterozygosity index, i.e. the probability that two alleles from a given locus, selected at random from two populations, will be different. So the higher the genetic distance between two populations, the longer the separation period between them and the larger the difference in vertical cultural characteristics. The second genetic dimension (*Genetic distance II*) is constructed using different theoretical properties, see Spolaore and Wacziarg (2009) (2009), for a more detailed description [16] Measures for ethnolinguistic distance are provided by Melitz and Toubal (2014b) and Egger and Toubal (2016). One such measure is based on the language tree classification and the other one is based on lexicostatistics. In more detail, the first ethnolinguistic variable (*Linguistic distance I*), is based on the categorization of the linguistic tree proposed by Lewis et al. (2009). The ethnologue database lists and describes around 7,000 languages, by continent and country. The second ethnolinguistic variable (*Linguistic distance II*) derives from the Automated Similarity Judgment Program (ASJP). This program evaluates the degree of lexical similarity between 200 words (or eventually 100 words) in a list that was first defined by Swadesh (1952). Because the construction of ethnolinguistic variables is based on two official languages per country, those measures might under-represent the cultural diversity within a country and might underestimate the cultural distance among countries. As an example, the ethnolinguistic diversity between Italy and Spain is equal to the one between Italy and Argentina or Italy and Mexico, given that the official language spoken in all countries is Spanish [17]. However this representation might not take into account the geographical relatedness and potential interactions between the countries. I have proposed and constructed new geo-weighted ethnolinguistic measures for this study:*Geo-linguistic distance I* and *Geo-linguistic distance II*, where the ethnolinguistic

[16] Economic contributions already exploit genetic differences to predict how cultural divergences affect economic behaviour. For example, Spolaore and Wacziarg (2009) demonstrate that the genetic distance has a statistically and economically significant effect in explaining cross country income differences, even after controlling for related measures of geographical distance, climatic differences, transportation costs, and historical, religious, and linguistic distance. More recently, Spolaore and Wacziarg (2014) provide evidence regarding the role of cultural distance on the diffusion of fertility decline across European countries. In parallel, Guiso et al. (2009) exploit genetic distance to study how persistence in cultural distance affects economic trade behaviour.

[17] For a more in-depth analysis of the languages associated to each country in the dataset refer to Table A.1 in Appendix A of Melitz and Toubal (2014b). The Table reports, for each country, the official language spoken, the main spoken language and the two relevant languages exploited for the construction of the ethnolinguistic distance metrics and their relative weights.

distance measures are weighted in respect of the relative geographical distances between countries [18].

1.3.2 Estimation strategy

The empirical analysis is based on a particular quasi-longitudinal dimension of data. The estimation exploits weekly variability in the duration of marriages to investigate the role of ethnic assortative mating on marital dissolution incidence [19]. The analysis is based on completed marriage durations for families who dissolved in the period between 1995 and 2012, while censored durations are derived for the rest of unions. The research reports the Kaplan-Meier non-parametric results for of the survivor function, graphically. For observed failure times $t_1, t_2, ..., t_j, ..., t_k$, the hazard rate θ_j at time t_j is defined as the conditional probability of observing the failure at time t_j, having survived up to that time: $P[T = t_j | T \geq t_j]$. The estimator of the hazard function is the ratio: $\hat{\theta}_j = \frac{d_j}{m_j}$, where d_j is the number of marriages that dissolves at time t_j and m_j is the number of marriages at risk of failure at the beginning of period t_j. Given the above notation the Kaplan-Meier estimator of the survival function correspond to the sample analogue:

$$\hat{S}(t) = \prod_{j|t_j \leq t} (1 - \hat{\theta}_j) = \prod_{j|t_j \leq t} \left(\frac{m_j - d_j}{m_j} \right) \tag{1.1}$$

Hence, the Kaplan-Meier estimate is the product over time of the conditional probability of surviving past time t_j given survival up to that time. It is estimated as the ratio of the number of individuals that experienced the event during the specific period, over the number of subjects at risk up to that period (i.e. all failures and censored observations at or after that time) [Kaplan and Meier (1958)] [20]. Consequently, in order to estimate the differential in the risk of dissolution among families, a Cox PH regression model was exploited [Cox (1972)].

[18] For the purpose of this analysis, cultural diversity within family is defined as the cultural distance between the countries of birth of spouses. The empirical estimations make use of the genetic distance measures available from Spolaore and Wacziarg (2009) and of the ethnolinguistic distance measures available from Melitz and Toubal (2014b). For a more accurate description of the variables construction refer to the above mentioned references. In addition geographical distance measures have been exploited. Results confirm the main conclusion presented.

[19] In principle, daily variability could be exploited to get rid of interval-censoring, i.e. when the completed duration is observed but only in interval form. Nevertheless, durations in economic applications are often interval-censored, under the assumption that the effect of interval-censoring is sufficiently small it can be ignored [Cameron and Trivedi (2005)].

[20] Right censoring implies that at the time of the observation, the transition event has not occurred, hence the total length of time between entry to and exit from the state is unknown. In presence of censoring, OLS estimates on spell duration might be biased both in case of exclusion of censored cases altogether from the estimation, because of potential selection issue, both imputing to censored observations a duration corresponding to the observed length as if they are complete, disproportionately weight under-recording higher failure time. Binary dependent regression model might allow to mitigate censoring bias explicitly modelling the transition probability to the event, although in this case cross individual variability in durations is not accounted for, reducing available information.

The research mainly modelled transitions rather than mean duration. It modelled the probability that a marriage that has survived up to time t will end at time $t+1$, i.e. the interest lies in modelling the conditional probability of the existing marriage ending, as a function of the marriage's duration, controlling for a set of observable covariates $\mathbf{x}' = (x_1, x_2, ..., x_n)$. The Cox PH model describes the conditional hazard rate as factored into two separate functions as follows:

$$\theta(t|\mathbf{x}) = \theta_0(t)\exp(\mathbf{x}\beta) \tag{1.2}$$

where $\theta_0(t)$ represents the nonparametric baseline hazard and it is a function of t alone. The baseline hazard is then augmented by the effect of covariates, which enter the model through the linear predictor: $\exp(\mathbf{x}\beta)$. Estimation is based on exact matching of marriages with respect to all relevant assortative mating dimensions: ethnicity, age of spouses, past marital history and educational attainment. Labor market controls are also included. The specification adds a dummy for prenuptial financial agreements (disjoint vs joint management of family wealth). The population size, of the husband and wife community, is included as a proxy for the interactions in the local reference marriage market. Finally regional and year fixed effects are included to account for heterogeneity among regions and over time. All controls refer to the time of marriage, which are predetermined at the moment of marriage. Selection on observables of course does not allow to achieve causal identification. For instance, the spouse's marriage selection might be affected by unobserved characteristics that contemporaneously affect marital stability.

Partial likelihood estimation is required to estimate the parameters of interest. Denoting $t_1, t_2, ...t_j, ..., t_k$, the observed failure times, the probability that spell j is the actual spell that ends equals the conditional probability of failure for spell j divided by the conditional probability that a spell of any individual in the risk set (i.e. the set of individuals who are at risk of failing just before the jth ordered failure: $R(t_j) = \{l : t_l \geq t_j\}$) fails, as:

$$\begin{aligned} P[T_j = t_j | t_l \geq t_j] &= \frac{P[T_j = t_j | T_j \geq t_j]}{\sum_{l:t_l \geq t_j} P[T_l = t_l | T_l \geq t_j]} \\ &= \frac{\theta_j(t_j|\mathbf{x}_j;\beta)}{\sum_{l:t_l \geq t_j} \theta_l(t_l|\mathbf{x}_l;\beta)} \end{aligned} \tag{1.3}$$

The partial likelihood function is the joint product of the above probabilities over the k ordered failure times [21]. The Cox PH model presents various advantages. First, com-

[21] In presence of ties the standard Peto-Breslow approximation [Peto (1972) and Breslow (1974)] is exploited and the partial likelihood function is adjusted consequently. Cox estimates reported are based on Breslow method for ties. As robustness also Efron approximation is exploited for handling tied values [Efron (1977)], results remain consistent.

pared to binary response models, it exploits variability in the marriage duration and not just in dissolution choices. More importantly it allows to control for censoring selection. Secondly, among all the survival estimation models, it uses a partial likelihood estimation, i.e. the baseline hazard is left unspecified. Considering that fully parametric models might produce inconsistent parameter estimates in case of misspecification of the density function, a semiparametric method, that requires less than complete distributional specification, is generally preferred and empirically exploited as standard method for survival analysis [Cameron and Trivedi (2005)]. The Cox regression model is a proportional-hazards model. It assumes that the hazard ratio is constant over time, i.e. the hazard rate depends on the ratio of covariates but does not depend on t. Graphical and goodness of fit tests were performed to confirm the proportional hazard assumption [22].

1.4 Marriage Formation: Empirical Evidence

1.4.1 Differentials in Assortative Mating Observables Dimensions

Table 1.1 gives illustrative statistics for the most relevant and observable assortative dimensions that were used as covariates, such as age, marital history and educational attainment,[Weiss and Willis (1997); Greenwood et al. (2014); Kalmijn et al. (2005); Frimmel et al. (2013); Bruze et al. (2015) and Choo (2015)]. Descriptives are also reported separately for all marriages and for marriages that ended in separation, and respectively for homogamous and heterogamous families. Approximately 7% of the unions ended in separation, during the studied time interval (1995-2012). In the sample, 86.8% of unions are among Italian natives, while heterogamous unions account for 11.2% and migrants for 2%. Of those unions, the percentage of dissolution is 6.8% in homogamous families and 7.8% in heterogamous families, respectively.

In terms of wife's age at the time of the marriage, the vast majority of unions were made when the wife was 26-30 years old and these also corresponded to the more unstable marriages. In general, the older the entry into marriage, the higher the stability of the marriage. This pattern is similar for homogamous and heterogamous marriages. However, in heterogamous marriages, the women were generally older at the time of the marriage; 10% were between 36-40 years old and 13% were more than 40 years old. In terms of the difference in ages between the spouses, in the vast majority of the unions, around 43.9%,

[22] Different tests provide mixed evidence. In particular, on the one hand the global test reject the null hypothesis of proportional hazard, this can be due to the burden of controls included in the model; indeed single coefficients test provide mixed results. On the other hand the fitting of the model proxy in a very precise way the observed survival probabilities derived from Kaplan-Meier non parametric estimates for different specifications [Garrett, 1997]. Finally, test of zero slope in a generalized linear regression of the scaled Schoenfeld residuals on time have been performed. This is equivalent to test that the log hazard-ratio function is constant over time. The result provides evidence in favour of the proportional hazard assumption.

marriages were formed between contemporary spouses (the age distance is 0-3 years) and in 17% of the unions the wife was older than husband. In case of heterogamous unions, only 28% had a spousal age difference of 0-3 years, while age differences of more than 7 years were common (34%). Hence for comparable age of the wife at the time of marriage, the assortativeness in terms of age is quite dissimilar for homogamous vis--vis heterogamous unions. This seems to translate in an increase in marital instability. Panel A in Figure 1.1 shows the age assortative mating distribution graphically, by family type. It can easily be seen that heterogamous marriages have a lower degree of age assortativeness compared to the other family types.

Where marital history is concerned, mixed marriages are more likely to be second or further marriages, both for the wife (a 5.7% gap versus homogeneous marriages) and the husband (a 5.1% gap versus homogeneous unions), or for both at the same time. Finally, no remarkable differences emerge in the educational assortative mating distribution among family types. Panel B of Figure 1.1 displays the pattern, graphically.

1.4.2 Differentials in Childbirth

Childbirth choices are particularly relevant in terms of the stability of a marriage [Becker et al. (1977), Weiss and Willis (1985) and Browning et al. (2014)]. Evidence about the incidence and timing of child bearing seems to support this. Where separations were concerned, Table 1.1 uncovers a significant difference between homogamous and heterogamous marriages, in the probability of their having children, at about 16.4% (p-value 0.000). The difference is less pronounced at the extensive margin, in terms of number of children per family (0.04 gap).

If, on one hand, we interpret children as a marital specific investment, ex-ante more stable families tend to have a higher rate of childbirth. On the other hand, the higher the rate of childbirth in a union, i.e. the higher the marital investment, the higher the stability of the union, because it increases the exit cost from the marriage. As a consequence, a differential childbirth might explain part of the differential in the marital dissolution of heterogamous as compared to homogamous, families. Heterogamous couples might postpone child bearing if they anticipate that the marriage could well fall apart. In doing so, they deprive the marriage of a bond, assuming that having children lessens the likelihood of marital dissolution [23].

In addition, fertility choices might affect marital stability through another channel. Ac-

[23] To notice that those figures refer to marriages that experience a dissolution, hence in principle less stable. For instance, from vital statistics registries, the total fertility rate (TFR) of foreign women is equal to 2.37 in 2012, while the TFR of Italian women is of 1.29. Hence the observed difference between homogamous and heterogamous marriages in the probability of having children might well be related to anticipated differences in marital stability rather than differences in preferences for child bearing between native and foreign women. Further data are needed to deepen this conclusion.

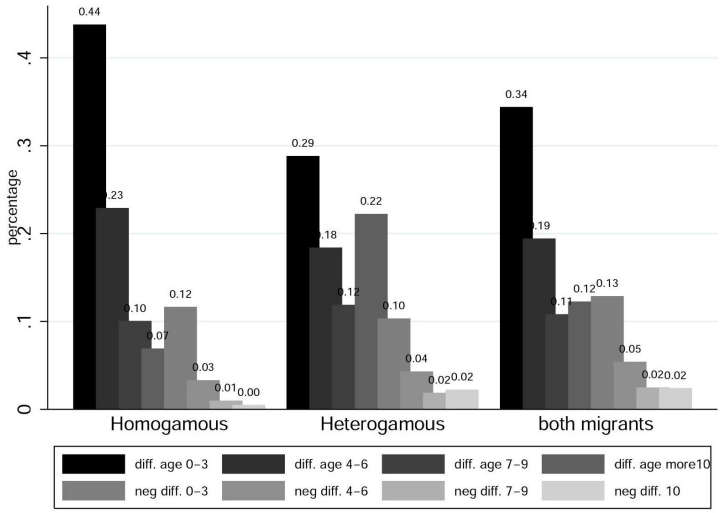
Table 1.1: Descriptive Statistics: Assortative Mating Dimensions

	Panel A: Sample of all Marriages			Panel B: Sample of Separated Marriages		
	All Marriages	Homogamous	Heterogamous	All Separations	Homogamous	Heterogamous
	(1)	(2)	(3)	(4)	(5)	(6)
Sample Share (%)		86.8	11.2		85.9	12.8
Age Assortative Mating						
Wife 16-20 age	0.038	0.039	0.036	0.065	0.065	0.068
Wife 21-25 age	0.226	0.227	0.208	0.309	0.311	0.297
Wife 26-30	0.376	0.386	0.307	0.379	0.39	0.312
Wife 31-35	0.207	0.207	0.208	0.158	0.156	0.168
Wife 36-40	0.079	0.075	0.108	0.052	0.048	0.079
Wife more 40	0.074	0.066	0.133	0.037	0.03	0.076
Same Age or Age Diff 0-3	0.419	0.439	0.288	0.401	0.425	0.25
Age Diff 4-6	0.223	0.229	0.184	0.222	0.231	0.173
Age Diff more 7	0.19	0.169	0.341	0.202	0.178	0.361
Age Diff 1-3	0.115	0.116	0.103	0.116	0.117	0.103
Age Diff more 4	0.053	0.047	0.084	0.059	0.049	0.113
Marital History Assortative						
Husband First -Wife First	0.888	0.907	0.76	0.891	0.912	0.763
Husband First -Wife Second	0.035	0.027	0.084	0.039	0.03	0.097
Husband Second -Wife First	0.047	0.041	0.092	0.046	0.038	0.092
Husband Second -Wife Second	0.03	0.025	0.064	0.024	0.02	0.048
Education Assortative Mating						
Husband Low - Wife Low	0.308	0.303	0.326	0.33	0.325	0.365
Husband Low - Wife Medium	0.152	0.151	0.164	0.167	0.165	0.183
Husband Low - Wife High	0.019	0.017	0.037	0.014	0.012	0.031
Husband Medium - Wife Low	0.064	0.062	0.076	0.073	0.071	0.088
Husband Medium - Wife Medium	0.282	0.29	0.23	0.281	0.291	0.214
Husband Medium - Wife High	0.049	0.049	0.049	0.038	0.038	0.035
Husband High - Wife Low	0.008	0.008	0.013	0.007	0.006	0.014
Husband High - Wife Medium	0.046	0.046	0.044	0.042	0.042	0.038
Husband High - Wife High	0.072	0.074	0.061	0.048	0.05	0.032
Disjoint	0.551	0.552	0.582	0.575	0.566	0.648
Husband Profession	1.673	1.689	1.612	1.733	1.755	1.621
Wife Profession	1.39	1.389	1.41	1.396	1.401	1.378
Fertility dimensions						
Prob. of children				0.508	0.53	0.365
Number of children (Couple)				1.379	1.383	1.343
Number of children (HH)				1.439	1.455	1.353
Observations	4464429	3872958	501663	306647	263563	39133

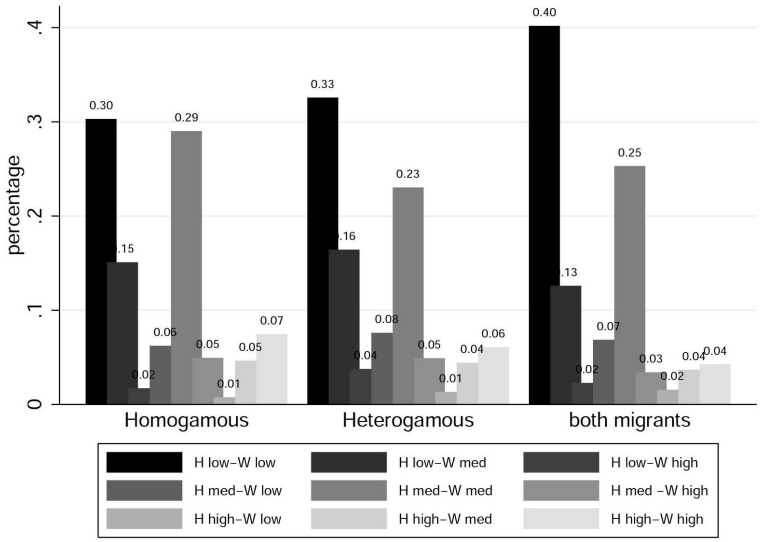
Note: Marriage and separations ISTAT datasets (1995-2012), Italy. ADELE elaboration. All assortative mating variables are dummies. Table reports averages by marital status and family type: Panel A reports averages for all sample of marriages, Panel B report averages only for separated unions. Ethnic Assortative mating: *Homogamous* are unions among Italian spouses; *Heterogamous* refers to families where one spouse is a native Italian and the other is a migrant; *Migrants* refers to unions of two migrants. Age assortative mating includes six variables for the wife class ages (16-20, 21-25, 26-30, 31-35, 36-40, more 40) and five dummies for spousal age distance. Marital history assortative includes four (4) dummies for the matching of spouses in first or second and further marriages. Education assortative mating includes nine (9) dummies for matching of spouses' educational level (the classification are: low for compulsory education, medium for undergraduate, i.e. high school degree and high for graduate and post-graduate education). Disjoint is a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Labor market controls* include professional status (blue-collar, white-collar, director, self-employed and entrepreneur) of both spouses. childbirth dimensions are available only of separated families: descriptives reports the probability of having children by family type; the number of children per union (couple), as children born from that union before and after marriage; and number of children within leaving within the household (HH), which includes also children born from different unions but residing within the family.

Figure 1.1: Assortative Mating in Age and Education by Family Type (1995-2012)

Panel A: Age Assortative Mating



Panel B: Education Assortative Mating



Source: Marriage ISTAT datasets (1995-2012), Italy. ADELE elaboration. Panel A reports the age assortative mating distribution by family type. Age distance of spouses is evaluated as the difference between husband and wife age at marriage, negative distance correspond to unions with wife older than husband. Panel B reports the education assortative mating distribution by family type. Educational attainment included 3 categories: low education for compulsory education level, medium for undergraduate and high for graduate and post-graduate education.

cording to the cultural transmission literature, the choice to enter into homogamous marriages is affected by inter-generational cultural transmission preferences [Bisin and Verdier (2000) and Bisin and Verdier (2011)]. Preferences for the vertical transmission of cultural values to children, represent an incentive for ethnically homogamous marriages [24]. In a similar way, the same preferences for children's socialization, might induce parents, in heterogamous families, to dissolve the marital union and to employ a single socialization tool whenever the distance between the spouses' culture weakens the effectiveness of the direct socialization process. In respect of this, Dohmen et al. (2012) have showed that single mothers' socialization technologies are equivalent to those of mothers in homogamous families and significantly more effective than heterogamous mothers' technologies, relaxing theoretical model assumptions as in Bisin and Verdier (2000). Indeed, Dohmen et al. (2012) suggest that mating randomly and then remaining a single mother could be interpreted as a potential alternative strategy to positively assortative marriage, under the assumption that the probability of the father's gaining custody is lower. This could mean that the higher the mother's probability of gaining child custody, the higher the incentive for dissolution. Evidence of a positive relationship between the likelihood of gaining child custody and marital dissolution gender differences might shed some light on this direction. Unfortunately, as mentioned earlier, childbirth information was only available for the separation and divorce datasets. Apart from the evidence, already discussed, the lack of data reduces the ability to deeply investigate this lead. The question remains open for future investigation.

1.5 Marriage Dissolution: Duration Analysis

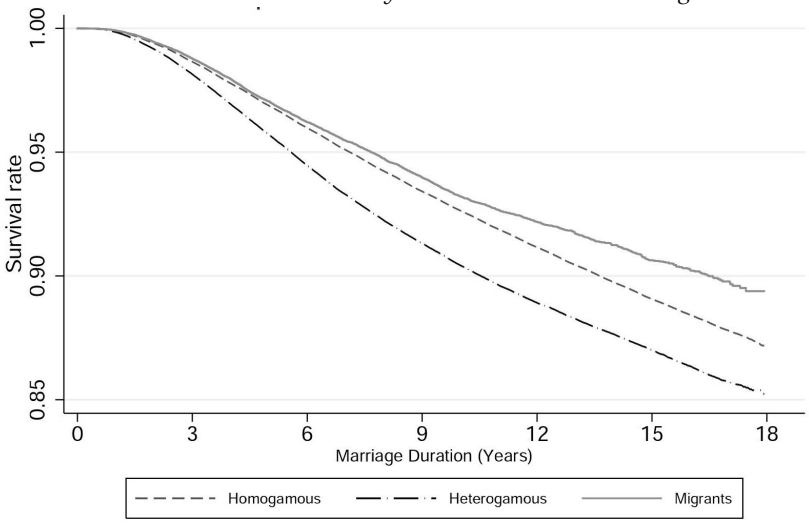
1.5.1 Non Parametric Survival estimates

At a first descriptive level, and focusing on families that experienced dissolution, the average duration of a marriage is 81 months, for homogamous families, while the average duration of the marriage for heterogamous families is 72 months. The difference is statistically significant - around 1 year (*p-value* 0.000). The incidence of separation for homogamous families is 10% while, for heterogamous families, it is higher at 13%. Non-parametric Kaplan-Meier estimates of the survival probability are reported in Figure 1.2. They show a differential in marital stability among family type, graphically. Both figures report the proportion of marriages that were not dissolved (on the vertical axis), after a stated num-

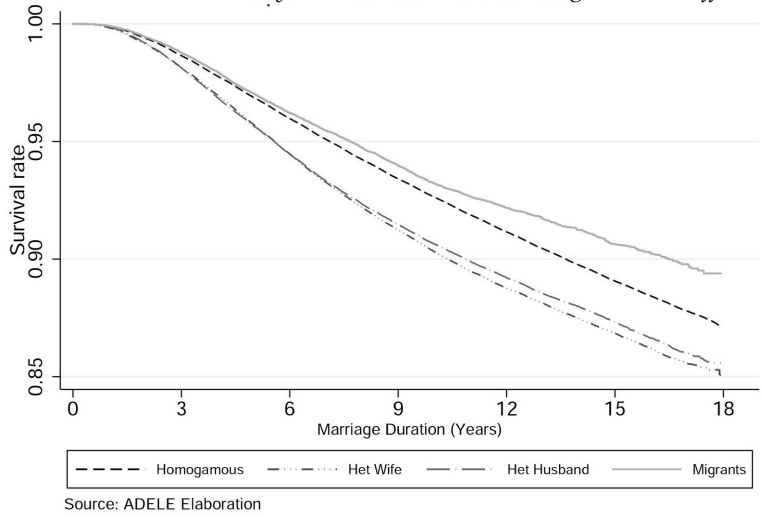
[24] Relative to this cultural transmission mechanisms, the literature demonstrates that the demand for cultural pluralism, which induces long-term persistence in cultural traits, arises from the interaction of rational individuals in culturally heterogeneous social environment [Bisin and Verdier (2000), Bisin and Verdier (2001)]. In particular, on the one hand, individuals respond to social heterogeneity through vertical socialization effort. On the other hand, agents select into homogamous marriages in order to have access to a more efficient vertical socialization technology [Bisin et al. (2004)].

Figure 1.2: Kaplan-Meier Survival Estimates by Family Types

Panel A: KM Estimates by Ethnic Assortative Mating



Panel B: KM Estimates by Ethnic Assortative Mating- Gender differences



Source: Marriage and Separation ISTAT datasets (1995-2012), Italy. ADELE elaboration. In Figure, Panel A reports the non-parametric Kaplan-Meier estimates of the survival probability to separation by family type. *Homogamous* refers to unions among Italian native spouses; *Heterogamous* refers to unions between one Italian native and one migrant; *Migrants* refers to unions among migrant spouses. Figure, Panel B reports the non-parametric Kaplan-Meier estimates of the survival probability to separation by family type, conditioning for the gender of the foreign spouse in heterogamous marriages.

ber of years (indicated on the horizontal axis). The calendar dimension is ignored in these representations, i.e. differences in the start date of the marriages plays no role in the non-parametric estimation. Evidence shows a differential in the survival probabilities among families starting from 5 years after marriage. The estimated marital survival probability is systematically lower in heterogamous couples, for all the durations of marriages considered in the study, compared to the homogamous ones. Indeed, a positive and significant distance exists in the marriage survival functions [25]. Conversely, migrants' families always seem to be more stable. The log-rank test statistic provides evidence to reject the null hypothesis of equality of the survival functions in the three groups (*p-value* 0.000) [26]. In particular, 91.7% of homogamous marriages are still intact after 10 years, while the same figure is 89.5% for heterogamous marriages and 92.6% for migrants. After 15 years of marriage, percentages drop respectively to 88.7%, 86.6% and 90.4%. Results point to an increase in the unconditional probability of marital dissolution due to a reduction in assortative mating along the ethnicity dimension. Moreover a significant difference in marital stability is evidenced among heterogamous families, depending on the foreign spouse gender, i.e. heterogamous families with a non-Italian wife are observed to be more stable than heterogamous families with a native wife and a non-Italian husband, see Panel B of Figure 1.2.

1.5.2 *Cox PH Regression Model Estimates*

Ethnic Assortative Mating

A Cox PH regression model is exploited to estimate the differential in the likelihood of separation among family types. Coefficients and hazard rates are reported in 1.2. Column (1) reports the estimated unconditional differences, in the likelihood of separation of heterogamous and migrants' families, compared to homogamous baseline probability of separation. Year and region fixed effects are included, to account for differences in time trends and regional heterogeneity, as concerns the differential socio-cultural perception of the separation phenomenon. The results underline the fact that differences in ethnic assortative mating, between spouses, increases the risk of separation. In particular, when compared to Italian native spouses, heterogamous families (with a foreign husband or wife) show a 21.1% higher risk of separation. Conversely, migrant families show a lower probability of separation of 24.3%. Multivariate Cox regression estimates are reported in Columns (2-7).

[25] Confidence intervals are evaluated from Greenwood estimate of the variance of the survival function:

$$\hat{V}[\hat{S}(t)] = \hat{S}(t)^2 \sum_{j|t_j \leq t} \frac{d_j}{m_j(m_j - d_j)} \tag{1.4}$$

[26] The Log-rank test is performed both unconditionally and by stratified sample, including as controls the covariates described above.

Table 1.2: Risk of Marital Dissolution: Ethnic Assortative Mating Differences

	Share (%)	Baseline Estimates (1995-2012)			Sub-duration Estimates (1995-2000)	(2001-2006)	(2007-2012)	Citizenship
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic Assortative Mating								
Homogamous (reference group)								
Heterogamous	11.24	.191***	.146***	.136***	.091***	.182***	.163***	.250***
		[1.211]	[1.157]	[1.146]	[1.096]	[1.199]	[1.177]	[1.284]
		(.005)	(.006)	(.006)	(.008)	(.009)	(.019)	(.007)
Migrants	2.01	-.278***	-0.409***	-.375***	-.409***	-.342***	-.411***	-.530***
		[0.757]	[0.664]	[0.687]	[0.665]	[0.710]	[0.663]	[0.589]
		(.016)	(.016)	(.020)	(0.030)	(0.022)	(0.042)	(.021)
Observations		4464429	4464429	4464429	1674292	1479605	1310532	4464429
Assortative Mating controls		No	Yes	Yes	Yes	Yes	Yes	Yes
Labour market controls		No	No	Yes	Yes	Yes	Yes	Yes
Husband's community pop		No	No	Yes	Yes	Yes	Yes	Yes
Wife's community pop		No	No	Yes	Yes	Yes	Yes	Yes
Region FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (1995-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Family type shares refers to the overall sample. The reference omitted group is of *Homogamous* Italian couples. *Heterogamous* refers to families where one spouse is native and the other one is migrant. *Migrants* is a dummy equal to 1 for families where both spouses are migrant. Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. Columns (4-6) report estimates for specific marriage duration categories. Column (7) exploits citizenship classification of migrants to define family types. Robust standard errors in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

In Column (2) relevant observable assortative mating dimensions are included. The differential in the risk of separation related to ethnic background decreases slightly, suggesting a positive correlation among assortative mating dimensions. Nevertheless, the gap driven by ethnic distance, remains positive and statistically significant for both heterogamous and migrants families at around 15.77% and 33.6%, respectively, with opposite signs. Column (3) includes the full set of controls, adding in labor market controls and husband and wife community population size [27]. Column (4-6) report sub-period estimates. Results suggest a non-monotonic pattern of marital instability over time. It increases in the period 2001-2006 and slightly decreases afterwards, both for heterogamous and migrants unions. Finally, the results presented in Column (7) use citizenship classification. As a further confirmation of previous conclusions, heterogamous families exhibit a higher risk of separation, compared to marriages among Italians, at around 28.4%. As mentioned earlier, the citizenship classification might overestimate the role of endogamy in explaining marital dissolution. Those individuals that acquire Italian citizenship through naturalization are more likely to be integrated into the host country, hence cultural differences are reduced. Moreover, because of Italian citizenship acquisition regulations, individuals born abroad, of Italian parents, are entitled to Italian citizenship. This implies that their inclusion in the native group overestimates the distance with respect to heterogamous families, assuming that they have more easily assimilated natives' customs, even if they grew up in different socio-cultural environment.

In summary, ethnic differences between spouses positively affects the risk of separation. Results support the heterogamous hypothesis [Becker (1973) and Becker et al. (1977)]. Dissimilarities in the inputs that enter as complements in the household production function (such as ethnicity), comparably, reduces the expected gains of mixed marriages and increases the risk of marital dissolution. Point estimates are consistent in magnitude with previous empirical evidence as reported by Zhang and Van Hook (2009) and Frimmel et al. (2013) for different countries, suggesting that the phenomenon is not country specific. On the other hand, the higher stability of the migrants' group is particularly interesting. Considering the low rate of separation of Italy, the evidence might suggest that even if migrants experience higher dissolution rates in their country of origin, they are relatively stable compared to natives. A strong family bond might be a result of difficulties in assimilation. Therefore, in absence of a social network support, the domestic network becomes essential, thus strengthening family cohesion.

[27] Labour market choices within family are of course equilibrium outcomes of the family optimization problem of production and consumption and hence endogenous. Including labour market dimensions at the moment of marriage mitigates potential endogeneity bias, even if anticipatory effects might still be in place. Nevertheless, labour market outcomes allows to proxy for the income and wealth status of the family, as a relevant control for marital stability. Results show that the inclusion of such covariates slightly modifies the differential in risk of separation, but point estimates remain coherent with previous analysis.

Parallel conclusions might be extended to all other observable spouses matching characteristics. In Table A.2, the following considerations emerge. Marital stability monotonically increases with spouses age. For instance, compared to a marriage where the wife's age is between 16 and 20 years, the risk of separation decreases by 35.5% when the wife enters the marriage aged between 21 and 25 years. The same difference increases to 82.5% for wives who marry at more than 40 years.Lehrer (2008) refers to this negative relation as a maturity effect. He underlines how the process of information acquisition evolves with age, in respect of both spouses' own marital preferences (self-acknowledgement effect), but also because of marriage market selection mechanisms (the learning effect). Moreover, in terms of assortativeness, the higher the age difference among spouses, the higher the risk of separation. For comparable age differences, in particular, whenever the wife is older than the husband the risk of marital dissolution increases. The marriage gains dynamic, in respect to age assortativeness, is discussed by Choo (2015). Where marital history is concerned, second or further marriages show a higher instability with respect to first marriages. The result parallels Bruze et al. (2015)'s results. They also demonstrate that individuals, entering into second marriages, have a weaker unobserved preference for marriage compared to those entering a first marriage. Hence, the higher instability of second and further marriages might be a consequence of selection mechanisms and is not due to the fact that second marriages are inherently less productive. Consistent evidence is also reported in Table A.3, Columns (2-3). Once again, similarly to Bruze et al. (2015), the results underline a clear asymmetry between the two genders. Second marriages for women are 63% more unstable compared to first marriages, while second marriages for men show a 33% higher risk of separation. A potential explanation for this asymmetry is that women bring children from previous marriages into their new marriages, because they more likely obtain the custody of the children after dissolution. Finally, educational attainment has a positive effect on marital stability. At the same time however, the higher the educational distance between the spouses the higher the risk of marital dissolution, as in Bruze et al. (2015) and Siow (2015).

Ethnic Assortative Mating: Gender Gap

Table 1.3 shows the comparable estimates that come from analyzing heterogamous families, by the gender of the foreign spouse, separately. Partialling out the effects of covariates, the probability of separation is 18.9% higher in heterogamous families with a non-Italian wife compared to homogamous Italian families, while the equivalent difference for heterogamous families with a non-Italian husband is 12.5%. Interestingly the comparison suggests the presence of a significant gender differential in marital ethnic composition, of

Table 1.3: Risk of Marital Dissolution: Ethnic Assortative Mating Differences - Gender Gap

	Share (%)	Baseline Estimates (1995-2012)			Sub-duration Estimates (1995-2000)	(2001-2006)	(2007-2012)	Citizenship
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic Assortative Mating: Gender gap								
Homogamous (reference group)								
Heterogamous Wife	7.79	.181***	.134***	.117***	.107***	.149***	.013	.196***
		[1.199]	[1.144]	[1.125]	[1.113]	[1.160]	[1.013]	[1.216]
		(.007)	(.007)	(.007)	(0.010)	(0.010)	(0.024)	(.008)
Heterogamous Husband	3.45	.212***	.169***	.173***	.063***	.263***	.434**	.414***
		[1.236]	[1.184]	[1.189]	[1.065]	[1.301]	[1.543]	[1.513]
		(.009)	(.009)	(.009)	(0.013)	(0.015)	(0.028)	(.012)
Observations		4464429	4464429	4464429	1674292	1479605	1310532	4464429
Assortative Mating controls		No	Yes	Yes	Yes	Yes	Yes	Yes
Labour market controls		No	No	Yes	Yes	Yes	Yes	Yes
Husband's community pop		No	No	Yes	Yes	Yes	Yes	Yes
Wife's community pop		No	No	Yes	Yes	Yes	Yes	Yes
Region FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (1995-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Family type shares refers to the overall sample. The reference omitted group is of *Homogamous* Italian couples. *Heterogamous* families are split with respect to the gender of the foreign spouse: *Heterogamous Wife* are families where the wife is born abroad Italy, reversely *Heterogamous Husband* are families where the husband is the foreign spouse. *Migrants* estimates are not reported in the Table. Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. Columns (4-6) report estimates for specific marriage duration categories. Column (7) exploits citizenship classification of migrants to define family types. Robust standard errors in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

about 6.4%. These results mirror Frimmel et al. (2013)'s results [28].

Differences in the cultural composition of mixed unions, across gender, are reported in Figure 1.3. The graph shows the percentage of heterogamous marriages by macro-area of origin and the gender of the foreign spouse. For both genders, the vast majority of marriages are celebrated with spouses from EU countries. But while Italian men are, largely, matched with Central-East European and South-America women, Italian women are more likely to marry North-African men. The distributions of the country of origin, by gender of the foreign-spouses is reported in Figure A.6, for the most represented origins. Asymmetries in marital distribution across macro-areas reflect asymmetries in cultural distances.

Ethnic Assortative Mating: Across Generation Assimilation

As mentioned above, due to the regulations concerning Italian citizenship acquisition and using information on citizenship acquisition (Italian born, naturalized Italian and non-Italian), it is possible to classify migrants with respect to their generations' order in first and second or further generation migrants. This classification leads to five different family categories: heterogamous first, heterogamous second, migrants first, migrants second and migrants mixed [29]. Heterogamous first marriages present a higher risk of separation at 15.2% compared to Italian homogeneous marriages. More interestingly, heterogamous unions of second generation migrants appear as stable as homogeneous unions. Analyzed together, results suggest that migrants are increasingly assimilated to Italian natives. For instance, the higher the generation order of the migrants, the less important the role played by the spouses' ethnic differences in explaining marital instability. The reported result is a small piece of evidence concerning the process of assimilation, but of major importance given the fast growing rate of second generations over time.

For instance, the percentage of second generation children from the total number of children, born in Italy in 1995, equals 1.7%. The same percentage increases to 9.4% in ten years (2005), and increases by nearly ten times again by 2012 reaching 15%. Once again,

[28] Looking at the spouse who submit the demand of dissolution, gender differences emerge. In general, for all family types, the vast majority of couples submit a join demand of marital dissolution (both for separations and divorce proceedings). Looking instead at disjoint submissions, in the most part of the cases dissolution is driven by wife unilateral preferences. For instance, among *Heterogamous Husband* families, the difference between the percentage of disjoint proceedings initiated by wife with respect to husband is of 19.8%, while the equivalent difference in *Heterogamous Wife* families is of 4%, suggesting that native-spouses in Heterogamous unions are more likely to unilaterally exit marriage. Even more interesting, looking at divorces, husband decisions overcome wife ones in *Heterogamous Wife* families, where the percentage of disjoint proceedings initiated by husband is much higher with respect to wife (20.3% vs 11.9%). From a descriptive point of view, there seems to be evidence of a gender asymmetries in unilateral dissolution decisions in mixed unions, with a relative prerogative in favour of native-spouses.

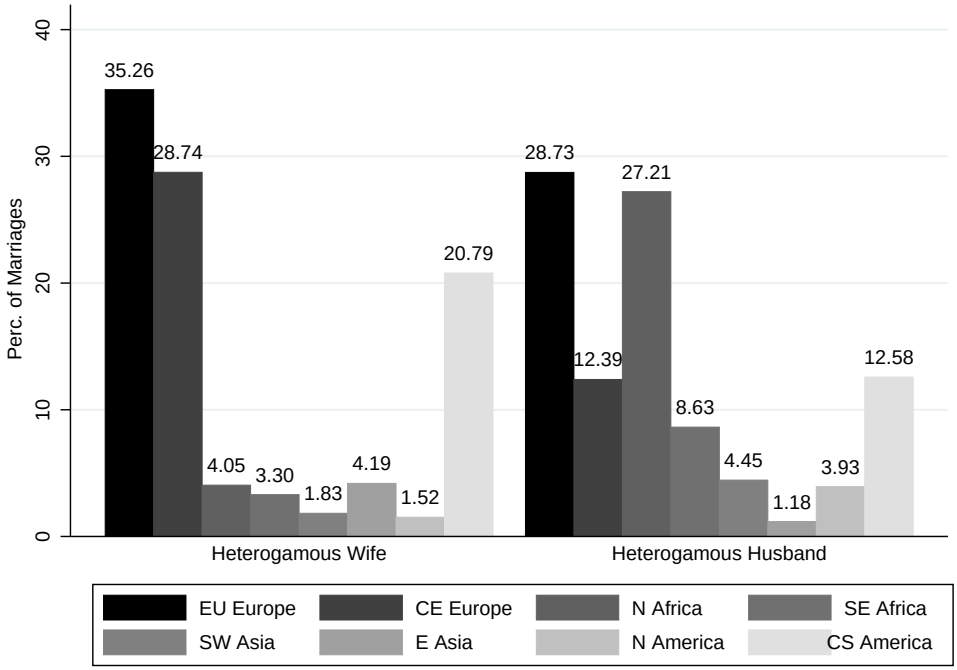
[29] Given data availability it's not possible to discriminate between second and third or further generations. This misclassification is not a major concern, since the migration phenomenon in Italy was particularly recent at the end of nineties. Estimates might be interpreted as a lower bound.

Table 1.4: Risk of Marital Dissolution: Ethnic Assortative Mating by Generation Order

	Share (%)	Baseline Estimates (1995-2012)			Sub-duration Estimates (1995-2000)	(2001-2006)	(2007-2012)
		(1)	(2)	(3)	(4)	(5)	(6)
Ethnic Assortative Mating by Generation Order							
Homogamous (reference group)							
Heterogamous First	11.04	.199**	.150***	.142***	.107***	.176***	.171**
		[1.220]	[1.162]	[1.152]	[1.113]	[1.192]	[1.186]
		(.006)	(.006)	(.006)	(.008)	(.009)	(.015)
Heterogamous Second	0.21	.028	.057*	.013	-.0,812***	.257***	-.499
		[1.021]	[1.059]	[1.013]	[0.443]	[1.293]	[0.607]
		(.025)	(.025)	(.025)	(.072)	(.028)	(.186)
Migrants First	2.93	-.339***	-.486***	-.444***	-.384***	-.487***	-.417***
		[0.712]	[0.615]	[0.642]	[0.681]	[0.615]	[0.659]
		(.018)	(.018)	(.018)	(.030)	(.025)	(.042)
Migrants Second	0.03	.186***	.272***	.202***	-.293**	.278***	.428
		[1.204]	[1.313]	[1.224]	[0.745]	[1.321]	[1.534]
		(.056)	(.057)	(.057)	(.077)	(.058)	(.317)
Migrants Mix Generation	0.05	.070	.032	-.003	-.839***	.191**	-.634
		[1.072]	[1.032]	[0.997]	[0.432]	[1.211]	[0.530]
		(.058)	(.058)	(.058)	(.192)	(.062)	(.352)
Observations		4464429	4464429	4464429	1674292	1479605	1310532
Assortative mating controls		No	Yes	Yes	Yes	Yes	Yes
Labour market controls		No	No	Yes	Yes	Yes	Yes
Husband's community population		No	No	Yes	Yes	Yes	Yes
Wife's community population		No	No	Yes	Yes	Yes	Yes
Region FE		Yes	Yes	Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (1995-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Family type shares refers to the overall sample. The reference omitted group is of *Homogamous* Italian couples. *Heterogamous* families are split with respect to the generation order of the migrant spouse: *Heterogamous First* refers to ethnic mixed families with a first generation migrant spouse while *Heterogamous Second* stands for mixed Italian-Foreign families with a second generation migrant spouse; *Migrants First* refers to migrants families of both first generation spouses while *Migrants Second* stands for families of both second generation spouses. Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. Columns (4-6) report estimates for specific marriage duration categories. Robust standard errors in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure 1.3: Percentage of Heterogamous Marriages by Macro-Area of Origin and Gender of the Foreign Spouse (Italy, 2005-2013)



Source: Marriages. Italy (2005-2013). Percentage of heterogamous marriages by Macro-Area of origin and gender of the foreign spouse. Percentage of computed over total number of marriages formed in the same period (2008-2014) in Italy. Macro-area Classification: EU Europe, European Union (25 countries); CE Europe: Center and East Europe ; N Africa: North Africa; SE Africa: South East Africa; SW Asia: South West Asia; E Asia: East Asia; N America: North America; CS America: Center South America.

the figures highlight the relevance and the transformation of the migration phenomenon in Italy.

Ethnic Assortative Mating: Cultural Distances

Evidence confirms the existence of a difference in marital dissolution probability between heterogamous and homogamous unions, where family types are defined based on the place of birth of the spouses. The richness of the data allowed the research to go beyond that result and to investigate whether heterogeneity among mixed unions, in terms of the cultural distance between the spouses, plays a role in explaining marital separation choices. In particular, I used the variability in the country of origin of foreign spouses across heterogamous families, to impute the distance between spouses' cultural values, proxied by the cultural distance of the birth country. The information on the country of origin of migrants is available from 2005 to 2012. The investigation explored two strategies. First, heterogamous families were divided into seven (7) different subgroups depending on the

macro area of origin of the foreign spouse [30]. Secondly, the cultural distance between the countries of origin of the spouses is used to measure their cultural distance. Results are robust to different measures of cultural distance.

Concerning the first strategy, Table 1.5 reports the Cox regression estimates. As before, the estimation includes all observable marital matching dimensions (age, education and marital history), as well as labor market controls and regional and year fixed effects. From Column (3), we see that no significant differences emerge among homogeneous Italian families and mixed ones if the foreign spouse comes from Europe, North America or Oceania. Significant differences can be found in all other cases. In particular, mixed unions with foreign spouses from Eastern Europe have an 11.3% higher risk of separation than unions among Italians. The corresponding difference increases to 42.5% for unions with spouses from South America and, basically, doubles when the migrants come from Africa. To get a sense of the ranking of those differences, the first two columns of Table 1.5 report the cultural distance of the macro area with respect to Italy, evaluated as the weighted average of cultural distances of countries of origin. There seems to be a close connection between the cultural background of spouses and the risk of marital dissolution, i.e. the higher the distance between the spouses' respective ethnic identities, the higher the risk of separation. Overall, evidence support the *heterogamous hypothesis*. Nonetheless using heterogeneity of origin, the importance of cultural-ethnic differences is a key driver in describing the differentials in marital dissolution. In order to investigate whether the family's cultural diversity is an important predictor of marital instability more fully, different measures of cultural distance were collected. An accurate description and their construction was presented in Section 3.5. For ease of comparability, the estimates use standardized cultural distance measures. Any correlation between cultural distance measures is described in Table A.5. In agreement with Spolaore and Wacziarg (2016), the correlation matrix underlines a high positive correlation among the different measures of cultural distance. They concluded that genetic distance is a summary statistic for differences in a wide range of inter-generationally transmitted human traits [31]. Table 1.6 reports the coefficient estimates of a Cox PH model on spouses' cultural distances. Overall, multivariate Cox regression coefficients indicate that cultural differences between spouses explain choices about marital dissolution, i.e. ceteris paribus, the higher the distance in the spouses' ethnic origins, the greater the risk of separation. All coefficients are actually positive and statistically significant. For instance, an increase of one standard deviation, in genetic distance, leads to an

[30] The list of countries that belong to each macroarea is reported in Table A.4

[31] They also provide evidence of a positive correlation between genetic and religious distance measures [Mecham, Fearon and Laitin (2006), and World Christian Database (2007)] and between genetic distance and distance in norms, values and attitudes considering all values-related questions appearing in the World Values Survey 1981-2010 Integrated Questionnaire.

increase in the risk of separation of 5.1%-5.6%, while an increase of one standard deviation, in ethnolinguistic distance, leads to an increase of about 8% in the risk of separation. Ethnolinguistic measures were weighted by the geographical relatedness of countries and report consistent results; very close to genetic distance point estimates, as expected. Lastly, I investigated the presence of non-linearity of the cultural distance effects on marital stability. The results are shown in Table A.6 of Appendix A.2. They underline the fact that, overall, an increase of one standard deviation in the spouses' cultural differences induces an increase in the risk of separation from 15% to 19%, with a diminishing rate for greater cultural differences. For example, where higher values of cultural dissimilarity are present, the marginal effect of an increase in cultural distance is lower compared to smaller values. In summary, heterogamous families show a higher risk of marital dissolution compared to homogamous marriages and this is greater again as the cultural distance between the spouses increases.

1.6 Conclusions

Nowadays, there is a growing body of literature, in family economics, that investigates who marries whom and why (Choo and Siow (2006b)), by looking at the mechanisms of family formation and matching in the marriage market, both theoretically and empirically. Several reasons motivate this increasing interest. Marital assortativeness is relevant to our understanding of intrahousehold interactions, allocation decisions and production technologies. It allows us to investigate changes in households that are brought about by economic, cultural and legal changes, as well as revealing about marital dissolution expectations. This paper explored the latter dimension: marital instability and aimed to contribute to the body of literature on that same perspective. In particular, I focused on one specific matching dimension, notably the cultural-ethnic identity of spouses. I investigated what role endogamy (i.e. marrying within the same ethnic group) played in characterizing the processes of marriage formation and dissolution.

I argued that the spouse's cultural-ethnic values are complement inputs within the household's production function. As a consequence, greater cultural distances between spouses induce higher uncertainty about marital outcomes, which increases the likelihood of dissolution. In order to test the above prediction, the empirical investigation exploit the quasi-longitudinal dimensions of the data. I combined information about marriages and dissolutions from different sources that covered the universe of the marriages celebrated in Italy from 1995 to 2012. Having access to actual data on marital dissolution, the research estimates the role of ethnic assortative mating on marital instability risk using a Cox PH duration model.

I presented evidence in favor of this heterogamous hypothesis. I showed a higher separation risk among interethnic families, compared to co-ethnic ones, matching marriages

with respect to all the relevant assortative mating dimensions, i.e. age, past marital history, education and labor market choices. Results suggest, as well, the existence of a foreign spouse gender effect, the separation risk is 6.4% higher in families with a non-Italian husband compared to heterogamous families with a non-Italian wife. This might be explained by the estimated difference between native and foreign mothers when it come to the probability of their being assigned child custody. When heterogeneity, in the generation order of migrants was examined, the evidence underlined a general assimilation of the migrants into the host country, in terms of decisions concerning family dissolution. Finally, the richness of data allowed the research to take a step forward, to investigate whether heterogeneity amongst mixed unions, in terms of the spouses' cultural distance, played a role in explaining marital separation choices. Evidence showed that the genetic and ethnolinguistic measures of cultural distance were informative about the incidence of marital dissolution, i.e. a standard deviation increase, in the cultural diversity of spouses, induced a 5-8% increase in the risk of marital dissolution [Spolaore and Wacziarg (2009); Melitz and Toubal (2014b) and Spolaore and Wacziarg (2016)]. Overall, evidence supported the *heterogamous hypothesis*. Nonetheless exploiting origin heterogeneity, it can be concluded that the magnitude of ethnic differences is a crucial element in describing marital dissolution variability.

When interpreting the results, I principally discussed intrahousehold channels and considered the differentials in preferences and technologies. I found that cultural diversity directly affects marital stability, increasing the uncertainty about marital outcomes. At the same time the increase in marital outcome variability lowers spouses' investment choices, in anticipation of a higher potential risk of dissolution. That aside, there are different mechanisms that might explain the same pattern. Differences in instability, across family types, might be explained by a carry-over of preferences towards divorce in migrants, in light of the low Italian separation rate. I chose to exclude this possibility. Firstly, this interpretation is inconsistent when the higher stability shown by migrants as compared to homogamous families. Secondly, I explored the correlation between the actual separation rate of migrants in Italy and the divorce rate in their country of origin, from 1970 and 1985, to determine if it could be viewed as a predictor of acquired socio-cultural attitudes to divorce. Interestingly, the estimates uncovered a mild negative correlation, that is, socio-cultural attitudes to divorce of the country of origin seems to not significantly explain marital instability in the host country [32]. Moreover, considering a search and matching framework, the differences in search costs and meeting rate for migrants might depict the same differential pattern in instability. Indeed, in the presence of friction in the marriage market, migrants face comparably higher search costs to find culturally similar spouses and they may also experience lower meeting rates for the same (similarly they face larger segregation costs, for example).

[32] Results not reported are available upon request.

Both premises imply a reduction in the reservation utility, leading in equilibrium to reach a lower expected matching quality. Disentangle preference versus opportunity incentives is a relevant point for future investigation.

Table 1.5: Risk of Marital Dissolution: Ethnic Assortative Mating by Country of Origin Macroarea

	Cultural distance		Share (%)	Duration Estimates (2005-2012)		
	Genetic I	Linguistic I		(1)	(2)	(3)
Ethnic Assortative Mating by Country of Origin Macroarea						
Homogamous (reference group)						
Heterogamous - Europe	22.25	5.89	4.51	-.014	-.090	-.090
	(18.02)	(.64)		[0.986]	[0.914]	[0.914]
				(.023)	(.023)	(.023)
Heterogamous - East Europe	20.82	5.97	3.26	.192***	.117***	.107***
	(11.16)	(1.04)		[1.212]	[1.124]	[1.113]
				(.024)	(.025)	(.025)
Heterogamous - Africa	138.56	6.89	1.36	.865***	.692***	.662***
	(141.75)	(.37)		[2.377]	[1.998]	[1.940]
				(.026)	(.026)	(.026)
Heterogamous - Asia	155.45	6.89	0.60	.153**	.116**	.085
	(81.67)	(.31)		[1.166]	[1.123]	[1.089]
				(.055)	(.055)	(.055)
Heterogamous - North America	71.53	5.38	0.51	-.044	-.041	-.041
	(31.25)	(.81)		[0.956]	[0.960]	[0.959]
				(.069)	(.070)	(.070)
Heterogamous - South America	104.91	3.75	2.16	.486***	.383***	.354***
	(71.47)	(.43)		[1.627]	[1.467]	[1.425]
				(.024)	(.024)	(.024)
Heterogamous - Oceania	26.15	6.10	0.08	-.123	-.069	-.080
	(2.27)	(.067)		[0.884]	[0.934]	[0.923]
				(.172)	(.174)	(.174)
Migrants			3.01	-.272***	-.509***	-.454***
				[0.762]	[0.601]	[0.635]
				(.030)	(.031)	(.031)
Obs.				1743834	1743834	1743834
Assortative Mating controls				No	Yes	Yes
Labour market controls				No	No	Yes
Husband's community population				No	No	Yes
Wife's community population				No	No	Yes
Region FE				Yes	Yes	Yes
Year FE				Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (2005-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Family type shares refers to the overall sample. The reference omitted group is of *Homogamous* Italian couples. *Heterogamous* families are split in 7 different subgroups depending on the area of origin of the foreign spouse.Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. The first two Columns report the cultural distance of the macroarea with respect to Italy, evaluated as the weighted average of cultural distances (genetic adn linguistic distance proxies) of countries belongings to the area. Robust standard errors in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.6: Risk of Marital Dissolution: Ethnic Assortative Mating by Cultural Distance Proxies

	Genetic Distance I (Fst) (1)	Genetic Distance II (Nei) (2)	Linguistic Distance I (Ling. Tree) (3)	Linguistic Distance II (ASJP) (4)	Geo-Linguistic Distance I (Weight) (5)	Geo-Linguistic Distance II (CES Weight) (6)
Mean	40.08	7.34	.43	.66	.078	.13
	(176.33)	(34.28)	(1.26)	(1.86)	(.33)	(.48)
Estimates (sd point)	.055***	.050***	.077***	.070***	.052***	.055***
	[1.056]	[1.051]	[1.080]	[1.073]	[1.053]	[1.056]
	(.003)	(.003)	(.004)	(.004)	(.003)	(.004)
Observations	1743834	1743834	1743834	1743834	1743834	1743834
Assortative mating controls	Yes	Yes	Yes	Yes	Yes	Yes
Labour market controls	Yes	Yes	Yes	Yes	Yes	Yes
Husband's community population	Yes	Yes	Yes	Yes	Yes	Yes
Wife's community population	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (2005-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Each column corresponds to a different estimation. Cultural distance is measured in terms of genetic and ethnolinguistic distance. The genetic distance metrics (*Genetic distance I* and *Genetic distance II*) are constructed on the *coancestor coefficients*: indices of heterozygosity, i.e. the probability that two alleles at a given locus selected at random from two populations will be different (Cavalli-Sforza and Piazza, 1994) (*Source*: Spolaore and Wacziarg (2009)). The ethnolinguistic distance measures are based on the language tree classification and on lexicostatistics analysis. The first ethnolinguistic variable (*Linguistic distance I*) is based on the linguistic tree categorization proposed by Lewis et al. (2009). The second ethnolinguistic variable (*Linguistic distance II*) derive from the Automated Similarity Judgment Program (ASJP) (*Source*: Melitz and Toubal (2014b) and Egger and Toubal (2016)). To account for geographical relatedness and potential interactions, new weighted ethnolinguistic variables have been constructed: *Geo-linguistic distance I* and *Geo-linguistic distance II* where the ethnolinguistic distances are weighted with respect to the relative geographical distance between the countries. For a more accurate description of the variables construction refer to the above mentioned references. All variables are standardized. Table reports mean and standard deviation of origin variables. Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. Robust standard errors are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Chapter 2

A Study of Marriage, Fertility and Divorce: Cultural-Ethnic Socialization of Migrants in Italy [1]

Abstract

In this paper we present a structural model of marital matching and intra-household decisions and estimate its parameters on Italian data. We consider a transferable utility marriage matching model along the ethnic dimension. Parents choose fertility, investments in the cultural socialization of children, and possibly divorce. The aim of the paper is to study how cultural socialization affects intra-household decisions; specifically, marital formation, fertility, and marital dissolution. Theoretically, we show that differences in transmission technologies have implications for marital choices, fertility and divorce patterns. We estimate model parameters, exploiting administrative data on marriages, separations and births, from Italy (1995-2012). Socialization frequencies come from the Condition and Social Integration of Foreign Nationals Survey (2011-2012). Intermarriage, fertility, separation and socialization rates are consistent with strong preferences of parents toward cultural socialization of children to their own ethnic identity, proxied by language transmission.

Keywords: Intermarriage, Marital Matching, Cultural Transmission, Migrants Integration.
JEL Classification: D1, J12, J13, J15.

[1] This chapter is joint with Alberto Bisin (New York University).

2.1 Introduction

In this paper, we investigate marital matching along the cultural-ethnic identity of the spouses. We aim at identifying the role cultural-ethnic traits play in marriage choices and at uncovering the implications of marital sorting along cultural-ethnic identities on consequent intra-household decisions. In particular, we provide an explanation of several stylized facts in marriage markets, e.g., the strong positive assortative mating along cultural-ethnic identities and the relative lack of stability of intermarriages, which centers around the role of cultural socialization of children. To this end we propose a structural model to study the process of family formation and intra-household decision making in a context where ethnic differences between spouses potentially matter both in terms of preferences and technologies for household production. We estimate model parameters on Italian data, exploiting the variability in ethnic identity across immigrants. The empirical analysis rests on rich administrative individual level data.

The main argument of the present research is that the process of cultural socialization of children is relevant in explaining the differential in household investments and stability across family types. We investigate the relevance of this novel channel, both theoretically and empirically. To illustrate it, we consider a transferable utility marriage matching model along the cultural-ethnic dimension of spouses, and we nest within this matching framework, a collective household model, in which parents choose fertility, investments in the cultural socialization of children, and possibly divorce.

We build upon the seminal contribution of Choo and Siow (2006b), considering a transferable utility (TU) marital matching model in a two-sided market without frictions. We model the joint marital utility as the sum of a systematic component, that only depends on observable spouses characteristics, and an idiosyncratic component reflecting individual unobserved heterogeneity. This idiosyncratic component is assumed to be additively separable into a wife-and a husband-specific components, which only depend on the observable characteristic of the potential spouse, avoiding any possible interaction in unobserved tastes. For given distributional assumptions, the structure of the model allows to translate the matching equilibrium conditions into a series of discrete choice problems. As a consequence, preferences are described by an additive random utility model. The joint marital utility includes a systematic component, related to the cultural socialization process of children as well as fertility and divorce choices, while individuals-specific idiosyncratic components aim to capture non-monetary returns from marriage.

On top of that, we outline an intra-household collective model. Within marriage, parents choose fertility, investments in the cultural socialization of children, and possibly divorce. By affecting household production technologies, cultural differences between spouses matter. Indeed, families care about socializing their children and are endowed with tech-

nologies to transmit their own ethnic trait to their children. We interpret the cultural transmission process as the interaction of two forces: the direct socialization effort of parents within the family and the indirect influence of the society at large, in line with anthropological literature on cultural transmission [Cavalli-Sforza (1981) and Boyd (1988)] and economic contributions [Bisin and Verdier (2000) and Bisin and Verdier (2001)]. In particular, extending Bisin and Verdier (2000)'s model, we assume that the direct socialization probability results from the interaction of individual parents' efforts, defined in a cooperative setting. Coordinated or conflicting socialization incentives in homogamous and heterogamous couples, respectively, imply different vertical technologies of cultural transmission. For instance, families where parents share the same cultural traits have access to a more efficient socialization technology in the transmission of their cultural traits as compared to families where parents share different cultural traits. Moreover we allow socialization technologies to depend on marital status. The expected marital values are thus affected by parents' endogenous investments in the cultural socialization of children, as well as by fertility and divorce choices. Hence, individuals match in the marriage market anticipating that marital choices will affect their future marital value via different socialization technologies.

We show that differences in transmission technologies have implications for fertility and divorce patterns and we show that the socialization channel implies complementarity in cultural traits of spouses, from a theoretical perspective. Moreover the model explain asymmetric household behaviours of minority ethnic groups as opposed to the majoritarian one.

Our empirical analysis focuses on interethnic marriages, exploiting variability in ethnic identity across migrants within Italy, in accordance with the continuing increase of migratory phenomena. Indeed, Italy experiences different waves of immigration with heterogeneity both in time and ethnic composition. At the beginning of 2014, the number of foreign residents registered in Italy was more than 4.9 million and they accounted for 8.1% of the total resident population. This is in comparison to 2013, when the same percentage was around 7.4% and as little as ten years earlier, in 2003, it was 2.6%. Migratory inflows have affected the Italian marriage market: the percentage of interethnic marriages nearly tripled during the period of investigation from 1995 to 2012.

We take advantage of rich administrative data, at the individual level, on marriages, births and separations, from Italy. The data covers the universe of marriages formed in Italy from 1995 to 2012, and the universe of births and separations registered in Italy in the same time period. We have constructed a unique quasi-longitudinal dataset, matching marriage records with birth and separation records, through time invariant observable spouses characteristics. The longitudinal structure of data provides a dynamic representation of intrahousehold decisions, starting from the moment of marital formation to potential child-

birth and dissolution decisions. We proxy the cultural-ethnic transmission by the language spoken at home by parents with children, relying on the observation that the ethnicity and spoken language are relevant related culturally specific attributes and both allow the direct transmission of cultural characteristic across generations. Socialization data come from the Condition and Social Integration of Foreign Nationals Survey (2011-2012). In line with theoretical predictions, the observed intermarriage, fertility, separation and socialization rates are consistent with strong preferences of parents toward cultural socialization of children to their own ethnic identity, proxied by language transmission. In particular we document the existence of strong positive assortative mating preferences along ethnic lines, homogamy rates are particularly high with respect to random matching. Moreover, marital preferences are heterogeneous both across ethnic-groups and markets and we uncover significant asymmetries between the majoritarian Italian group and the minority ethnic groups. In addition, we underline the existence of asymmetries between homogamous and heterogamous families in childbirth investments and divorce choices. Finally, by looking at the Italian language socialization probability, we document a differential in transmission technologies between family types and marital status. For instance, homogamous minorities families have a strong preference toward the transmission of their own language, hence the probabilities of speaking Italian at home are lower compared to those in heterogamous families. Observed differences in socialization technologies between marital states are in line with theoretical arguments.

We discuss an estimation strategy of structural model parameters and their identification. The main parameters of interest in the model are the cultural relative intolerance parameters, representing the relative distance in values of a parent in having a child sharing is own cultural trait rather than acquiring a specific different trait. They mean to capture the strength of parents cultural transmission desire. We estimate model parameters via a minimum distance estimator. We consider the empirical frequencies of gains to marriage (derived from the marital matching function), fertility and separations rates and socialization frequencies. We extend the matching model to a multi-market framework, exploiting geographical variability across provinces. Identification relies on exogenous variability in population vectors across markets and hinges on parameters' restrictions.

Our attention on cultural-ethnic lines and on the process of ethnic socialization is motivated by different reasons. First, a large body of literature have documented a strong resilience in ethnic traits, proxied by the country of origin [f.e. Fernández and Fogli (2009), Fernández (2013) and Figlio et al. (2016)] and the continuing increase of migration into Western countries makes ethnic boundaries more and more salient, as compared for example to religious ones [Dalla Zuanna (2008) and Kalmijn et al. (2005)]. Moreover, in line with above mentioned studies, we proxy the ethnic identity by the country of origin, that is objectively measured and not self-reported as for religious affiliation, where differences

between religious belonging and religiosity are of major importance. Finally, the language of the country of origin is a dimension that parents care to transmit to their children [Dustmann (1997), Ginsburgh and Weber (2011) and Chiswick and Miller (2014)].

The chapter is organized as follows. Section 3.2 provides a review of the literature. Section 3.4 outlines the theoretical framework, discussing the general setup, the timing of the model and choice components. Section 2.4 provides a discussion of theoretical results. Section 3.5 describes the data used in the empirical analysis and interesting patterns, while Section 2.6 presents the structural model, the identification and the empirical estimation strategy. Finally, Section 3.7 concludes and briefly presents the future directions of the investigation.

2.2 Literature Review

This paper builds on and bridges two literatures: the family economics literature, on the one hand, and the cultural transmission literature, on the other hand.

Starting from Becker's seminal contributions [1973, 1974, 1991] [2], family economics research developed following two main directions. On the one hand, it has generally focused on *who marries whom and why*, looking deeply into the assignment process of spouses, when competition for potential partners in the market affects individual choices [Choo and Siow (2006b) and Dupuy and Galichon (2014)] [3]. In particular, some studies focus on sorting mechanisms along age and deepen how demographic changes, as the baby boom generation or the gender differential in the mortality rate, affect the marriage market and the related matching process [Choo and Siow (2006b); Choo and Siow (2006a) and Choo (2015)]. Chiappori et al. (2012) mean to estimate the marital trade off between income and body mass index, considering a bidimensional matching model. Finally, a large literature analyses the educational assortativeness to examine whether and how spouses asymmetries in returns to education within marriage (i.e. marital college-plus premium) might be responsible for gender differences in high-education investments [Chiappori et al. (2009) and Chiappori, Salanié, and Weiss (Chiappori et al.)]. On the other hand, family economics research have studied how to characterize the households' decision making process and specifically how to model spouses' interactions [Lundberg and Pollak (1993); Chiappori et al. (2002); Blundell et al. (2007) and Del Boca et al. (2014)]. Deviating from the original unitary Becker perspective, these contributions have recognised that changes in prices or incomes that influence the family budget constraint might also affect the relative bargaining power of the partners.

[2] Becker (1973); Becker (1974). See also Becker et al. (1977) and Becker (1991).

[3] For a survey of the literature see also Browning et al. (2014) Chapter 7 and Chiappori and Salanie (2016).

Although the two strands of the literature have obvious mutual implications, they have grown separately. Compared to the existing literature, we develop our project in this direction. Indeed, we propose an intra-household decision model and we nest it into a transferable utility matching framework. On this, our analysis is in line with a recent paper of Chiappori, Salanié, and Weiss (Chiappori et al.), which emphasizes how these two strands of research are intrinsically related. Our research further contributes to bridging the current gap between cultural and family economic literature, with the primary objective of understanding better the reasons why cultural preferences affect individuals selection into the marriage market and its consequences on later intra-household decisions. Indeed, within the marital matching literature the cultural dimension has been greatly overlooked, despite the recognition of its relevance in explaining relevant economic outcomes and of the strong persistence of cultural traits across generations.

In the last decades, a growing literature investigates the relationship between economic outcomes and preference formation. A pioneering contribution to this literature is the paper of Becker and Mulligan (1997), which formalizes a model where individuals endogenously choose their own preferences, rather than those of their children. Recently, it is exactly this process of transmission of preferences, beliefs, and norms of behaviour the main object of study of several social sciences, with the final goal of rationalize the observed persistence over time of cultural identities.

Various social sciences interpret the cultural transmission process as the result of social interactions across and within generations [Cavalli-Sforza (1981) and Boyd (1988)]. The main economic insight to this stream of research reaches the conclusion that a demand for cultural pluralism arises from the interaction of rational agents in culturally heterogeneous social environment. Namely, the investment in socialization of the families responds to the exogenous incentives of the society at large. The fact that parents react strategically to their children social environment, having access to a direct costly socialization technology, is the essential element that allows the economic models to rationalize the sustained cultural heterogeneity of the societies in the long run, as opposed to melting pot predictions of a perfect convergence to a cultural homogeneous society [Bisin and Verdier (2000) and Bisin and Verdier (2001)]. On top of this, the same demand for cultural pluralism is internalized by individuals within marital decisions. Thus, individuals rationally select in homogamous with respect to heterogamous marriages in order to have access to a more efficient vertical socialization technology [Bisin et al. (2004)].

A large and growing body of research developed from this starting point. For example, Fernández et al. (2004) document the role of mothers labor choices in favouring the female labor participation of following cohorts. Doepke and Zilibotti (2008) and Klasing (2012) underline how the increase in economic power of the middle class during the British Industrial revolution contributes to explain the variation in risk attitudes, shaping the actual

economic growth. Fernández and Fogli (2006) and Fernández and Fogli (2009) provide evidence in favour of the persistence in fertility and work practices across cultures, as cultural and behavioural preferences of second and further generations migrants living in the US are predicted by cultural and behavioural attitudes present in the country of origin. Hauk and Saez-Marti (2002) and Tabellini (2008) focus on the effect of policy changes in institutions on current trust and social capital. Tabellini (2010) underlines the effect of the diffusion of specific cultural traits on economic development, through their impact on better functioning institutions. In a more recent analysis, Figlio et al. (2016) study the impact of long-term orientation differences on the educational achievement of immigrant students residing in the US, providing evidence that parents from long term oriented cultures are more likely to secure better educational opportunities for their children.

Starting from above findings and conclusions, our research expands the literature studying the implications of cultural socialization preferences on household formation, considering explicitly the competition in the marriage market, and later intra-household investment decisions.

2.3 Theoretical Framework

2.3.1 Setup

We construct a structural model of marital matching and intra-household decision making. Specifically, we assume a transferable utility (TU) marriage matching model along ethnic dimensions in a frictionless environment. In TU models, spouses implicitly transfer utility between each other in absence of transition costs. Transfers are endogenously determined as equilibrium outcomes, indeed they depend not only on the quality of the specific match but also on the set of available opportunities in the marriage market.

We assume a large marriage market, with a population of M men and F women. We postulates heterogeneity in cultural-ethnic identity of men and women. For simplicity, in this section we present the theoretical model under pure dichotomous cultural trait in the society, while we extend the econometric framework to multiple cultural-ethnic groups. Each man $m \in M$ and woman $f \in F$ has one cultural trait indexed by i and j respectively, where $i, j \in \{a, b\}$. As a consequence, in the marriage market there is heterogeneity in family types. Indeed within family, spouses might or might not share the same cultural traits. Compared to the unitary household model proposed by Bisin and Verdier (2000), we assume that each parent ($p \in \{m, f\}$) has the desire to transmit his own cultural values to children and has a socialization technology, which is increasing in his effort of socialization. Spouses interact cooperatively within marriage, which implies that intra-household decisions are Pareto efficient. We deviate from this assumption under the divorce case. For instance, following

a divorce, parents choose non-cooperatively their socialization efforts. As a consequence, parents in heterogamous families face conflicting incentives in the socialization of children, while parents in homogamous unions benefit from coordinate incentives, which implies an asymmetry in the vertical technologies of cultural transmission between family types.

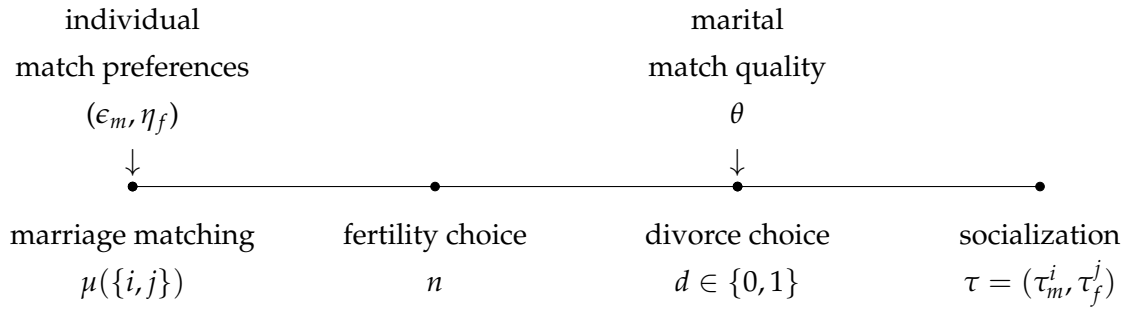
The expected marital gains are affected by parents' endogenous investments in the cultural socialization of children, and by fertility and divorce choices. Hence, individuals match in the marriage market anticipating that marital choices affect their future marital gains via different socialization technologies.

We describe below the timing and the characterization of preferences, in more details.

2.3.2 *Timing*

We propose a dynamic model of family formation and intra-household decision making. The timing of the model is represented below.

Figure 2.1: Timing of the model



As illustrated in Figure (2.1), at the first stage agents match in a frictionless marriage market with TU. In particular, we assume that the utility is transferable within marriage and upon divorce. We model the total marital utility as the sum of two components: a systematic component related to the cultural socialization process of children; and an idiosyncratic component, that captures non-monetary returns from marriage. In particular, we introduce an idiosyncratic structure at the level of individual matches, denoting by ϵ_m and η_f, men and women unobserved tastes, respectively. These random variables mean to capture individual specific deviations from the systematic component, unobserved from the econometric point of view. We assume that heterogeneous unobserved preferences are additive separable in preferences and depend only on observed cultural traits. This separability assumption excludes complementarity between unobserved spouses' characteristics, and allows to translate the two-sided matching problem to a series of one-sided

discrete choice problems. Hence, preferences are described by a random utility model[4]. When agents match in the marriage market, they anticipate the value of their future marital utility, forming expectations regarding their marital stability, represented by the random variable θ. θ is unknown at the moment of the marriage and it is revealed only successively. Marriages realize whenever the expected gains from the union are higher then the outside option of remaining single. For sake of simplicity, the utility from remain single is normalized to zero.

The marital union also gives spouses the possibility to have children. The utility that parents derive from children depends on their expected marital stability and on their expected socialization quality, as we will describe below. Afterwards, spouses update their knowledge on the quality of their marital union, observing the realization of θ. According to the realization of θ, individuals decide whether they remain married or to divorce. Unanticipated match quality shocks might trigger divorce [Weiss and Willis (1997), Chiappori et al. (2008)]. We consider θ as a match-specific random variable, ruling out any asymmetry between spouses within a marriage. We impose that its distribution is symmetric around the mean equal to zero, and it's independent across marriages. Uncertainty on the quality of the match explains divorce choices as rational deviations from ex-ante optimal decisions of entering the marital union, and represents how in real contexts the process of information acquisition within the marital relationship evolves over time.

At the moment of the divorce decision, spouses anticipate the differences in socialization technologies across marital states, where socialization investment decisions are endogenous in the model. It is convenient to describe the structure of preferences and technology while introducing the maximization problem of the agents. We proceed backwards, from the socialization component, to the divorce and the fertility components in turn. We conclude with the marriage optimal assignment.

We categorize marriages in terms of the ethnic types of the spouses. A marriage between a male of ethnic type $i \in \{a, b\}$ and a female of type $j \in \{a, b\}$ is denoted as a marriage of type $\{i, j\} \in T \{\{a, a\}, \{a, b\}, \{b, a\}, \{b, b\}\}$. Let τ_m^i and τ_f^j denote the socialization effort of, respectively, a male of ethnic type $i \in \{a, b\}$ and a female of type $j \in \{a, b\}$; and $\tau = (\tau_m^i, \tau_f^j)$. We interpret τ as the direct socialization probability, such as $\tau \in [0, 1]$ in all families. Let $d \in \{0, 1\}$ denote the family divorce status, where $d = 1$ denotes a marriage ended in divorce and $d = 0$ a continuing marriage. Let n denote the number of children in the marriage. Let finally q^i denote the fraction of agents in the relevant population of reference with ethnic trait $i \in \{a, b\}$, that is, the population distribution by ethnic trait.

[4]See subsection (2.3.6) for further distributional assumptions on idiosyncratic tastes.

2.3.3 *The socialization component*

Consider a marriage of type $\{i,j\}$ in divorce status d in a population distribution with a share q^i of type i. Both spouses socialize the children. We assume for simplicity all children turn out of the same trait. [5] The expected utility the male(of type i) obtains per child from the socialization process is denoted $W_m^i(\tau;\{i,j\},d,q^i)$, net of socialization costs. If we let $P_i(\tau;\{i,j\},d,q^i)$ denote the probability that the children are of type i and V_j^i the value to a father of type i of a child of type j, for i distinct from j [6], then

$$W_m^i(\tau;\{i,j\},d,q^i) = P_i(\tau;\{i,j\},d,q^i)V_i^i + (1-P_i(\tau;\{i,j\},d,q^i))V_j^i - c(\tau). \tag{2.1}$$

Assuming for simplicity that the mother is given custody of children in divorce, we posit socialization technologies, extending Bisin and Verdier (2000), as follows:

$$\begin{aligned}
&i=j \qquad P_i(\tau;\{i,j\},0,q^i) = \tau_m^i + \tau_f^j + (1-\tau_m^i-\tau_f^j)q^i \\
&i=j \qquad P_i(\tau;\{i,j\},1,q^i) = \tau_f^j + (1-\tau_f^j)q^i \\
&i\neq j \qquad P_i(\tau;\{i,j\},0,q^i) = \tau_m^i + (1-\tau_m^i-\tau_f^j)q^i \\
&i\neq j \qquad P_i(\tau;\{i,j\},1,q^i) = (1-\tau_f^j)q^i
\end{aligned} \tag{2.2}$$

According to the above specifications, the socialization technologies result from the interaction of both vertical and horizontal socialization forces. In a family of type $\{i,j\}$, both spouses choose their socialization efforts $\tau = (\tau_m^i, \tau_f^j)$ with which their preference parameters are direct transmitted to children. If the direct transmission of attitudes from parents to children fails, children type will be determined by the interaction with the society at large, since the likelihood of interaction with adults of each type depends on their representation in the society, i.e. q^i [7]. Restrictions on τ and q^i, together with functional form assumptions in (2.2), insure that socialization probabilities are bounded between 0 and 1.

Let $W(\tau;\{i,j\},d,q^i) = W_m^i(\tau;\{i,j\},d,q^i) + W_f^j(\tau;\{i,j\},d,q^i)$ represent the total utility from the socialization process, per child. Under TU, socialization efforts $\tau = (\tau_m^i, \tau_f^j)$ are chosen to maximize the total utility from the socialization process, net of socialization costs:

[5] In particular differences in socialization preferences regarding the sex and/or the birth order of children are left apart. Moreover, socialization externalities driven by spillover effects across siblings are not considered.

[6] We assume that $\Delta V^{ij} \equiv V_i^i - V_j^i > 0$, i.e. the difference in father preferences for having a child of the same type i rather than j is strictly positive and similarly for the mother of type j: $\Delta V^{ji} \equiv V_j^j - V_i^j > 0$.

[7] The assumption that cultural traits are picked at random refer to as unbiased oblique transmission. Other formulation based on positive or negative biases can in principal also be studied, yet they are rarely used in practice. See Boyd (1988) for an extensive analysis of negative, positive and frequently depend bias issue and Sáez-Martí and Sjögren (2008) for a theoretical discussion of their effect on long-run dynamics.

$$\max_{\tau} \; nW(\tau; \{i,j\}, d, q^i) \tag{2.3}$$

Under the assumption that socialization costs are quadratic, which we maintain in the analysis, the solution of this problem is independent of the number of children n. Let it be denoted

$$\tau\left(\{i,j\}, d, q^i\right) = \begin{pmatrix} \tau_m^i \\ \tau_f^j \end{pmatrix}\left(\{i,j\}, d, q^i\right)$$

2.3.4 *Divorce component*

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. Let

$$\tilde{W}(\{i,j\}, d, q^i) = W\left((\tau)\left(\{i,j\}, d, q^i\right); \{i,j\}, d, q^i\right).$$

Divorce will be optimally chosen, that is, the marriage will be rationally dissolved in application of the Becker-Coase Theorem, when

$$n\tilde{W}(\{i,j\}, 1, q^i) - n\tilde{W}(\{i,j\}, 0, q^i) > \theta$$

The divorce condition has a probabilistic translation. Given $F(\theta)$ the cumulative distribution of θ, we can then compute the probability of divorce:

$$Prob\left(d = 1 \mid \{i,j\}, n, q^i\right) = F_\theta\left(n\tilde{W}(\{i,j\}, 1, q^i) - n\tilde{W}(\{i,j\}, 0, q^i)\right).$$

The Law of Large Number guarantees then that a fraction $Prob\left(d = 1 \mid \{i,j\}, n, q^i\right)$ of marriages of type $\{i,j\}$ in a population distribution q^i will be dissolved in divorce. Let denote the divorce probability by $\pi\left(\{i,j\}, n, q^i\right) = Prob\left(d = 1 \mid \{i,j\}, n, q^i\right)$.

2.3.5 *Fertility component*

The marriage fertility choice will in turn be made optimally anticipating the probability of divorce. Let[8]

$$U\left(\left(\{i,j\}, n, q^i\right)\right) = \begin{bmatrix} \pi\left(\{i,j\}, n, q^i\right) n\tilde{W}(\{i,j\}, 1, q^i) + \\ \left(1 - \pi\left(\{i,j\}, n, q^i\right)\right)\left(n\tilde{W}(\{i,j\}, 0, q^i) + \theta\right) - c(n) \end{bmatrix}.$$

The quantity-quality trade off that characterizes endogenous fertility choices is captured

[8]For notational convenience, we omit to include individual specific marital preferences as they simplifies at this stage.

in the model [Becker (1960), Doepke (2015)]. The optimal number of children is determined by the expected socialization quality per child,driven by the likelihood of dissolution, and the marginal cost of raising them. Let $n(\{i,j\},q^i)$ denote the solution of the following problem:

$$\max_n \; U\left(\{i,j\},n,q^i\right) \tag{2.4}$$

2.3.6 *Marriage Matching*

Individuals sort in the marriage market along ethnic characteristics, anticipating their indirect marital utility under different potential matches. Let

$$\tilde{U}\left(\{i,j\},q^i\right) = U\left(\{i,j\},q^i,n(\{i,j\},q^i)\right).$$

Let $\mu\left(\{i,j\},q^i\right)$ denote the fraction of marriages of type $\{i,j\}$ in a population with distribution q^i, and $\mu\left(\{i,0\},q^i\right),\mu\left(\{0,j\},q^i\right)$ denote the fraction of unmatched men and women in the same population q^i, respectively.

The optimal stable assignment is the solution of the welfare maximization problem over all potential matches $(\mu\left(\{i,j\},q^i\right),\mu\left(\{i,0\},q^i\right),\mu\left(\{0,j\},q^i\right))$ [Shapley and Shubik (1971)] subject to the following feasibility and non-negativity constraints [9]:

$$\begin{aligned}
\sum_j \mu\left(\{i,j\},q^i\right) + \mu\left(\{i,0\},q^i\right) &= m_i \quad \forall i \in \{a,b\},\\
\sum_i \mu\left(\{i,j\},q^i\right) + \mu\left(\{0,j\},q^i\right) &= f_j \quad \forall j \in \{a,b\},\\
\mu\left(\{i,j\},q^i\right) \geq 0, \mu\left(\{i,0\},q^i\right) &\geq 0, \mu\left(\{0,j\},q^i\right) \geq 0.
\end{aligned} \tag{2.5}$$

In order to derive the marital matching function, we follow Choo and Siow (2006b). We introduce a stochastic structure at the level of individual matches, such that spouses' preferences are described by a random utility model. The idiosyncratic components (ϵ_m,η_f) capture individual specific deviations from the systematic preferences. Idiosyncratic components are observed by individuals at the moment of the match, but are unobserved from the econometric point of view. We assume that heterogeneous unobserved preferences are additive separable in preferences and depend on observable characteristics only, hence we denote them by $(\epsilon_m^{ij},\eta_f^{ij})$. This separability assumption excludes complementarity between unobserved spouses' characteristics. We still allow for sorting on unobservable dimensions, but not simultaneously from both sides of the market. The above assumptions reduce the

[9] We characterize the welfare maximization problem and its dual version in Appendix (B.1).

two-sided matching problem to a series of one-sided discrete choice problems, which are linked by the adding up formula generating the total surplus [Galichon and Salanié (2015) and Chiappori, Salanié, and Weiss (Chiappori et al.)]. The discrete choice problem of man i and women j is thus:

$$\begin{aligned} U_m^i &= \max_{j \in \{a,b\}} \left\{ \tilde{U}_m \left(\{i,j\}, q^i \right) + \epsilon_m^{ij}; \epsilon_m^{i0} \right\}, \\ V_f^j &= \max_{i \in \{a,b\}} \left\{ \tilde{U}_f \left(\{i,j\}, q^i \right) + \eta_f^{ij}; \epsilon_f^{0j} \right\}, \end{aligned} \tag{2.6}$$

subject to:

$$\tilde{U} \left(\{i,j\}, q^i \right) = \tilde{U}_m \left(\{i,j\}, q^i \right) + \tilde{U}_f \left(\{i,j\}, q^i \right),$$

where $\tilde{U}_m \left(\{i,j\}, q^i \right)$ and $\tilde{U}_f \left(\{i,j\}, q^i \right)$ represent the indirect marital utility of man m and women f, respectively.

We follow Choo and Siow (2006b) assuming that (ϵ, η) are independent and identically distributed random variable with a type I extreme-value distribution (Gumbel). Distributional assumptions of unobserved random preferences, translate the additive random utility models presented above in the multinomial logit model [McFadden (1974)] [10]. Straightforward construction of matching probability $\mu \left(\{i,j\}, q^i \right)$ for a man i of marrying a woman j under TU and logistic shocks imply:

$$\mu \left(\{i,j\}, q^i \right) = \frac{\exp \tilde{U}_m \left(\{i,j\}, q^i \right)}{1 + \sum_j \exp \tilde{U}_m \left(\{i,j\}, q^i \right)}$$

and

$$\mu \left(\{i,0\}, q^i \right) = \frac{1}{1 + \sum_j \exp \tilde{U}_m \left(\{i,j\}, q^i \right)}.$$

The systematic marital utility of man m of trait i in marring women j, net of transfers, is identified by the log odd ratio formula, as:

$$\tilde{U}_m \left(\{i,j\}, q^i \right) = \ln \frac{\mu \left(\{i,j\}, q^i \right)}{\mu \left(\{i,0\}, q^i \right)}.$$

Similarly, we can identify the net systematic marital utility of women f of trait j in marring a man i, by the corresponding log odd-ratio formula. For given equilibrium transfers, the marriage market clearing condition implies that the demand of women j from men i is equivalent to the supply of women j for men of type i, for all men and women types, i, j. Hence, we identify the marital value function as the sum of spouses indirect utilities, by the

[10] We include a proof in Appendix B.1.

following ratio:

$$\tilde{U}\left(\{i,j\},q^i\right) = \ln \frac{\mu\left(\{i,j\},q^i\right)^2}{\mu\left(\{i,0\},q^i\right)\mu\left(\{0,j\},q^i\right)} \tag{2.7}$$

Let $\varphi(\{i,j\})$ denote the RHS of equation (2.7).

2.4 Results

Proposition 1 (Socialization Investments). In cultural heterogeneous societies, $q^i \in (0,1)$, for $i,j \in \{a,b\}$, in equilibrium:

i. in homogamous married families ($i = j$), the parents socialization efforts are strictly positive, $\tau(\{i,j\},0,q^i) > 0, \forall p \in \{m,f\}$ and the probability of successful socialization of children is greater than the horizontal socialization rate, $P_i(\{i,j\},0,q^i) > q^i$; the parents socialization efforts, $\tau(\{i,j\},0,q^i) > 0, \forall p \in \{m,f\}$ are monotonically increasing in cultural intolerance preferences, ΔV^{ij} and decreasing in the population distribution, q^i;

ii. in heterogamous married families ($i \neq j$), the socialization effort of the parent with higher cultural intolerance is strictly positive, while the other parent does not invest in socialization. Let's assume $\Delta V^{ji} > \Delta V^{ij}$, then $\tau_f^j(\{i,j\},0,q^i) > 0; \tau_m^i(\{i,j\},0,q^i) = 0$, the probability of successful socialization of children to trait j is greater than the horizontal socialization rate, $P_j(\{i,j\},0,q^i) > (1-q^i)$ while $P_i(\{i,j\},0,q^i)) < q^i$. Parent j socialization effort, $\tau_f^j(\{i,j\},0,q^i)$ is monotonically increasing in his own cultural intolerance preference, ΔV^{ji}; decreasing in his spouse cultural intolerance preference, ΔV^{ij}; and finally increasing in the population distribution q^i. Under perfect symmetry between parents' cultural intolerance preferences, the direct socialization technology of heterogamous families nullifies, $\tau(\{i,j\},0,q^i) = 0$ and socialization probabilities equate to horizontal socialization rates, $P_i(\{i,j\},0,q^i) = q^i$ and $P_j(\{i,j\},0,q^i) = (1-q^i)$;

iii. in homogamous divorced families ($i = j$), the mother socialization effort is strictly positive: $\tau_f^j(\{i,j\},1,q^i) > 0$. Divorced families underinvest in socialization compared to married ones, $\tau(\{i,j\},1,q^i) \leq \tau(\{i,j\},0,q^i)$;

iv. in heterogamous divorced families ($i \neq j$), the mother socialization effort is strictly positive: $\tau_f^j(\{i,j\},1,q^i) > 0$. Divorced families overinvest in socialization compared to married ones, $\tau(\{i,j\},1,q^i) \geq \tau(\{i,j\},0,q^i)$;

Parents make costly investments in order to socialize their children, both in homogamous and heterogamous families. Socialization investments in homogamous families ben-

efit from coordinate incentives. Conversely, a positive socialization investment in heterogamous families hinges on cultural intolerance asymmetries. Indeed, in case of perfect symmetry between parents preferences, the model goes back to the Bisin and Verdier (2000) initial assumption of zero vertical transmission.

Parents have the incentive to positively socialize their children, both under marriage and divorce states, if socialization is available [11]. Conditioning on being married, parents' socialization investments are Pareto efficient, resulting from marital cooperative behaviour. Conversely, under divorce, mothers' socialization investments are socially inefficient. Positive or negative externalities in socialization between homogamous and heterogamous married families, respectively, translate in underinvestment or overinvestment effort provision under divorce.

In addition, homogamous families hold a more efficient socialization technology, compared to heterogamous ones, $\tau(\{i,j\},0,q^i \mid i=j) > \tau(\{i,j\},0,q^i \mid i \neq j)$ and $P_i(\{i,j\},0,q^i \mid i=j) > P_i(\{i,j\},0,q^i \mid i \neq j) > q^i$ in case of marriage, while under divorce, both families are equal in their socialization technologies, $\tau_f^j(\{i,j\},1,q^i \mid i=j) = \tau_f^j(\{i,j\},1,q^i \mid i \neq j)$.

Interestingly, Proposition 1 outlines an asymmetry in household behaviours between the minority ethnic group as opposed to the majoritarian one, because parents socialization choices interact with their social environment of reference. In particular, the direct efforts of parents in the intergenerational transmission of cultural values substitute the horizontal socialization channel, which implies that the direct efforts are larger for the minority group, ceteris paribus. Finally, any increase in cultural intolerance of parent of type i, ΔV^{ij} translated into an higher socialization quality of children successfully socialized to trait i. This positively affects parent i socialization investment, $\tau_m^i(\{i,j\},d,q^i)$. Notice in particular that, while socialization efforts in homogamous families, $\{i,j\}$ with $i=j$, respond only to cultural intolerance preferences ΔV^{ij}, heterogamous families socialization efforts respond to both parents cultural preferences ΔV^{ij} and ΔV^{ji}, in opposite directions.

Proposition 2 (Divorce Choice). In cultural heterogeneous societies, $q^i \in (0,1)$, for $i,j \in \{a,b\}$ and assuming positive fertility $n>0$, at the equilibrium:

i. in homogamous families ($i=j$), the difference between the expected marital utility under divorce and marriage is strictly negative, $n\tilde{W}(\{i,j\},1,q^i) - n\tilde{W}(\{i,j\},0,q^i) < 0$;

ii. in heterogamous families ($i \neq j$), the difference between the expected marital utility under divorce and marriage is strictly negative, $n\tilde{W}(\{i,j\},1,q^i) - n\tilde{W}(\{i,j\},0,q^i) < 0$;

iii. for given realizations of marital stability shock,θ and assuming $\Delta V^{ji} > \Delta V^{ij}$, the di-

[11] Potential extensions to non-negative probability of father custody assignment after divorce, will not change the results in terms of family behaviour.

vorce probability of heterogamous families is higher compared to homogamous families: $\pi(\{i,j\},n,q^i \mid i \neq j) > \pi(\{i,j\},n,q^i \mid i = j)$;

iv. for given realizations of marital stability shock, θ, and assuming $\Delta V^{ij} > \Delta V^{ji}$, the divorce probability of heterogamous families is higher with respect to homogamous minorities families, $\pi(\{i,j\},n,q^i \mid i \neq j) > \pi(\{i,j\},n,q^i \mid i = j)$ if $q^i < 1/2$ and viceversa for $q^i > 1/2$,

Two different and opposed mechanisms contribute to explain divorce choices: the differential in socialization technologies across marital states and the free riding incentives driven by different attribution of socialization costs. Proposition 2 outlines that, for given socialization preferences and technologies, only negative surprises on marital quality trigger divorce decisions. Indeed, for a neutral realization, $\theta = 0$, the families remain married, because of the marital gains arising from the presence of public good. This implies that the loss in socialization technologies associated to divorce decisions is not counterbalanced by a comparable reduction in socialization costs, both for homogamous and heterogamous families. This is quite intuitive for homogamous families, as they largely lose from marital deviations to divorce, because of less efficient cultural transmission technologies. In heterogamous families, instead, the divorce guarantees a more efficient socialization technology to the mother at a cost of the father socialization. The composite effect is negative.

Moreover, for a given realization of marital quality shock, when the mothers have larger cultural intolerance preferences compared to the fathers, the divorce probability of heterogamous families is higher with respect to homogamous families, independently of the distribution of cultural traits in the population. Thus, 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 children custody attainment. This perspective is pointed out by Dohmen et al. (2012). Indeed, they make the point that children are more strongly socialized by a divorced mother rather than in heterogamous families, while there is no significant difference in socialization compared to homogamous families. Finally, Proposition 2.iv states that the distribution of cultural traits in the population matters in the comparison of divorce decisions among families, when in heterogamous families the fathers have larger cultural intolerance preferences compared to the mothers. The result comes from the fact that being i the majoritarian group, $q^i > 1/2$, the loss in socialization technologies of homogamous families is counterbalanced by the increase in the favourable horizontal socialization and by free riding incentives on socialization costs.

Proposition 3.(Comparative Statics on Divorce Choices). In cultural heterogeneous societies, $q^i \in (0,1)$, for $i,j \in \{a,b\}$, and for positive fertility $n > 0$, at equilibrium,

i. the divorce probabilities in homogamous and heterogamous families, $\pi(\{i,j\},n,q^i), \forall i,j$, are decreasing in the number of children n;

ii. the divorce probability in homogamous families ($i = j$), $\pi(\{i,j\},n,q^i \mid i = j)$, is decreasing in the level of cultural intolerance of parents ΔV^{ij} and it is increasing in the proportion of the society with similar cultural trait q^i;

iii. the divorce probability in heterogamous families ($i \neq j$), $\pi(\{i,j\},n,q^i \mid i \neq j)$, is decreasing in the level of cultural intolerance of parent with trait i, ΔV^{ij}, and it is decreasing in the proportion of the society with opposite cultural trait q^i;

Divorce choices depends on the socialization quality of children, but also on the quantity of the fertility investment. Interpreting children as a marital specific public good, Proposition 3.i uncovers a negative relationship between the probability of divorce of a family and the number of children in that family.

The probability of divorce rests on the distribution of cultural traits in the population. Other things equal, the marital instability of homogamous families belonging to the majoritarian group i is higher with respect to the instability of homogamous minority couples. Indeed, marital deviations bring a socialization risk that is lower the higher the proportion of the population with a similar cultural trait i, i.e. the more favourable the horizontal socialization channel q^i. A symmetric result holds for heterogamous families. Finally, any increase in cultural intolerance of parent of type i, ΔV^{ij} translates into lower rates of marital dissolution.

Proposition 4. In cultural heterogeneous societies, $q^i \in (0,1)$, for $i,j \in \{a,b\}$, at equilibrium:

i. the fertility rates in homogamous and heterogamous families, $n(\{i,j\},q^i)$, are strictly positive;

2.5 The Data: Marital Selection, Fertility, Separation and Socialization in Italy

From the econometric point the view, our interest lies in estimating the parameters of the marital value function, observing the marital matching patterns, as well as the fertility, separation and socialization frequencies. In this section, we first present the data we use in the estimation and some preliminary evidence corroborating theoretical model predictions.

Compared to the theoretical model presented in Section (3.4), two points need to be discussed.

First of all, we expand the cultural-ethnic identity set, allowing for K cultural traits in the society ($K > 2 : i, j, k = 1, .., K$). The cultural-ethnic identity of migrants is proxied by their country of origin [12]. In particular, we consider a finite number of ethnic groups. First, we distinguish between Italians, as the majoritarian group, and migrants. Secondly, we aggregate minorities by foreign countries of origin to define 7 cultural-ethnic groups, respectively: Europe 27 (European EU countries), Other Europe, North Africa, Other Africa, South-West Asia, East Asia and America. Our classification reflects both the prevalence of each ethnic-group in Italy and the relative cultural distance of countries with respect to Italy. The list of countries belonging to each ethnic-group and their prevalence is reported in Table C.1. Figure **??** graphically shows the correspondence between our ethnic-group classification and the cultural distance of countries with respect to Italy, proxied by genetic distance, ethnolinguistic and geo-linguistic distance measures [13]. Overall, Figure **??** documents the presence of some geographical clusters in cultural dissimilarity by continent. In particular, our classification parallel the heterogeneity in genetic distance within Africa, between North-Arabic countries and Sub-Saharan countries, as well as the within Asia divide between Middle-East countries and East countries.

Secondly, we assume that the province is the relevant geographical unit of reference for marital and intra-household decisions [14]. This has two main implications. First, we presume that each province corresponds to a specific marriage market. The multi-market framework allows us to identify preference parameters exploiting geographical across-markets variability in populations vectors, under the assumption that different markets share common characteristics and preference parameters. Second, the reference group that affects parents socialization investments is the province, i.e. parents choose their socialization efforts, and previously their fertility and marital dissolution, considering the population composition of the province where they reside. Let g index the province of residence. In this perspective, our multi-market framework is in line with that proposed by Chiappori,

[12] Despite the fact that the migration phenomena is particularly recent in Italy, we consider as migrants both first and second or further generations. The ethnic-identity of first generation migrants is proxied by their country of birth, while for second or further generations, their ethnic-identity is identified by the country of birth of parents or by the mother's country of birth in case of asymmetries.

[13] The genetic distance measures the degree of genealogical relatedness between two populations and it is associated to the time elapsed since two populations' last common ancestors. The metrics for genetic distance is based on the genetic tree classification, as described in Cavalli-Sforza and Piazza (1994). It is defined based on the *coancestry coefficients*: the heterozygosity index, i.e. the probability that two alleles from a given locus, selected at random from two populations, will be different. This implies that the higher the genetic distance between two populations, the longer the separation period between them and the larger the difference in vertical cultural characteristics. Data on genetic distance metrics are available thanks to Spolaore and Wacziarg (2009). Measures for ethnolinguistic distance are provided by Melitz and Toubal (2014b) and Egger and Toubal (2016), and are based on the language tree classification proposed by Lewis et al. (2009). The ethnologue database lists and describes around 7,000 languages, by continent and country.

[14] We consider the province of residence of spouses at the moment of marriage. Within Italy migration choices are not taken into account, coherently with the theoretical model where we do not model endogenous location choices.

Salanié, and Weiss (Chiappori et al.), which exploits variability across cohorts and gender in educational investment decisions.

2.5.1 The Data

The empirical analysis uses original administrative Italian data, at the individual level, covering the period 1995-2012. From the data, we recover the bivariate distribution of marriages by cultural-ethnic group of spouses by province; the population vectors by ethnic group and marital status for each province; the fertility rates by ethnic group of spouses by province; the separation rates by ethnic group of spouses by province; the socialization probabilities by ethnic group of spouses by province; and finally the population distribution by ethnic group and sex for each province. A more in depth discussion of available data and sample selection is reported in Appendix B.2, while a synthetic description of the variables of interest is provided in Table B.1.

The empirical estimation is based on a unique quasi-longitudinal dataset, that we constructed matching marriage records with birth and separation records. We exploit an exact matching procedure thanks to time invariant dimensions. The longitudinal structure of data has two main advantages. First, it allows to follow households over time, having a complete representation of intra-household dynamic decisions, starting from marital choices to subsequent potential fertility and dissolution choices. Secondly, the matching process allows us to fix a particular time period characterized by increasing migration inflows. The final sample consists of 4,151,528 marriages, that cover the 92.58% of the universe of marriages celebrated in Italy from 1995 to 2012. The 87.28% of marital unions are homogamous Italian marriages, while the remaining percentage refers to marriages that involve at least one foreign spouse. First marriages account for the 88.28% of the total sample. The comparison of the two marital distributions suggest that remarriage rates are not systematically different across spouses ethnic groups. In the sample the fertility rate corresponds to 69.56% with an average of 1.54 children per family. Of all marriages, the 7% end up in separation in the first years of the marital union.

We restrict our attention to legal marriages, while cohabitations are not included in our sample. Despite the cohabitation rate increased in the last decade, data availability is very limited and only rely on Census data, which are available every ten years. Implicitly, we interpret the differential between legal marriage and cohabitation choices, in light of the fact that marriages entail an additional degree of commitment, which is especially relevant for long-term investments such as children socialization [Lundberg and Pollak (1993) and Chiappori, Salanié, and Weiss (Chiappori et al.)].

We derive the population vectors by ethnic group, sex and marital status from individual Italian Census data of 2001 and 2011. We select only adult unmatched individuals

(of more than 18 years of age). Census data classify the marital status of an individual as: never married, at present married, separated *de facto*, legally separated, divorced or widowed. Because the model allows for endogenous divorce choices, we consider that an individual is unmatched in case she/he is never married, legally separated, divorced or widowed. We take into account potential measurement error concerns due to truncation of unmatched population vectors. That is, the observed unmatched men and women in 2011 might well marry in future years, which leads to an underestimation of marital gains. Our analysis might be hampered in presence of systematic differences across ethnic groups in marital rates over time, namely if marital rates are systematically higher for some specific ethnic groups as compared to others, over the period. To shed a light on this point, we compare the vectors of unmatched men and women in 2001 by ethnic-group with those in 2011. We notice that unmatched rates increase quite symmetrically for all ethnic groups. The overall Spearman rank correlation test is as high as 0.88, and equal to 0.57 and 0.98 for available adult male and adult female, in turn [15], suggesting that the ethnic-group rank order remains stable over the period, especially for women. In addition, following Chiappori, Salanié, and Weiss (Chiappori et al.), we restrict the set of unmatched individuals to unmatched men and women after their marriageable age, defined as the 90 percentile of the age at marriage distribution for men and women, respectively, in 2001. Gains to marriage, computer from equation (2.7), are reported in Table B.5.

Socialization data come from the *Condition and Social Integration of Foreign Nationals* Survey (2011-2012). The survey is targeted to foreign residents in Italy with the aim of detect essential information on their living conditions, behaviours, attitudes and opinions. We exclude from our analysis, respondents who are not married and families without children, at the time of the interview. The final sample consists of 17,512 individuals belonging to 4,996 families and the 18.59% of those families are either separated or divorced. The survey is intended to provide a comprehensive representation of the socio-cultural as well as economic integration of foreign residents. In particular, we focus on the language spoken at home by parents with children, to recover socialization frequencies by spouses' ethnic group.

Our interest for intergenerational language transmission is twofold. First, the linguistic socialization is a relevant cultural dimension for parents [Dustmann (1997), Ginsburgh and Weber (2011), Clots-Figueras and Masella (2013) and Fouka (2016)]. Secondly, it allows us to study the degree of convergence of migrants to the host socio-economic environment. Indeed, several studies uncover a positive association between the proficiency in the destination language and migrants socio-economic integration, favouring the educational achievement of lag-behind children during compulsory schools and fostering employment

[15]The Spearman rank correlation test corresponds to the Pearson correlation between the rank values of the variables considered. It assesses the monotonic relationship between variables, without imposing ant linear relationship.

and earning opportunities [Dustmann and Fabbri (2003) and Dustmann et al. (2010)]. We delve into this relationship in our data, by looking at the correlation between our measure of linguistic socialization and different measure of socio-cultural integration of children, as for example the language spoken with school mates or friends out of school, or the nationality of school mates and friends out of school. Table B.7 shows that our measure of linguistic socialization, by capturing the persistence in migrants' cultural-ethnic identity, is negatively correlated with the measures of socio-cultural integration.

Finally, we derive the population distribution by ethnic group and province for the time period 1995-2012 from municipality records on the movements of the foreign resident population. Population shares by ethnic group and province are calculated thanks to municipality data on the total resident population, aggregated at the province level. The maps in Figure B.3, display the geographical heterogeneity in the population distribution between marriage markets, for the overall migrant population (first map) and separately for all other ethnic groups considered in our analysis.

2.5.2 *Patterns in the Data*

Several contributions in the sociological and economic literature studied marital assortativeness along spouses cultural-ethnic identities, mainly within the US environment. They provide evidence in favour of positive assortative mating along cultural lines in terms of racial identities, but also in terms of spouses' ethnicities and linguistic lines [Fu and Heaton (2008); Fryer (2007) and Schwartz (2013)]. Schwartz (2013), in particular, underlines the parallel between ethnic and linguistic homogamy, where both ethnicity and spoken language are relevant culturally specific attributes and both allow the direct transmission of cultural characteristic across generations. By looking at the Italian marriage market, we document the existence of a strong preference for homogamy along cultural-ethnic lines. Figure 2.2 reports the observed homogamy rates by spouses' ethnic group. Except for the majoritarian group of Italians, for all other ethnic groups in the analysis, homogamy rates are significantly higher compared to random matching rates, as represented by the 45 degree line. Strong preferences for positive sorting along cultural lines are common to all marriage markets within Italy, with heterogeneity even considering markets with similar population distribution. Moreover, we observe that mating preferences are heterogeneous across ethnic-groups, as they are particularly high for Other Africa and East Asia minorities.

Does marital matching influence later intra-household decisions? By reporting the fertility rate by parents ethnic group, separately for homogamous and heterogamous families, Table 2.1 documents the presence of asymmetries in childbirth investments across families. Comparing homogamous and heterogamous families, we observe a significant differential in fertility rates at the intensive margin for all ethnic groups. For instance the probability of

having at least one child in a homogeneous Italian family is of 73.5% while, the equivalent probability in a family with at least one Italian spouse is of 41.8%. We report in Column (3) fertility estimates for heterogamous families excluding families involving at least one Italian spouse, in order to control for external social incentives or economic differential opportunities. Results remain consistent. A similar pattern is evidenced at the extensive margin, looking at the average number of children, conditional on having a positive fertility. Observed asymmetries in fertility rates across family types support the hypothesis that spouses cultural differences lower the investment in marital-specific capital, possibly in anticipation of higher marital instability.

We investigate marital dissolution, by looking at separation rates. Separation rates are quite low in Italy and especially so for homogamous marriages. We report separation rates by family type and ethnic group in Table 2.2. We observe a positive differential in marital instability of heterogamous marriages as compared to culturally homogeneous unions, in line with previous results [Becker et al. (1977); Kalmijn et al. (2005), Bratter and King (2008) and Zhang and Van Hook (2009)]. For instance, the probability of marital dissolution in homogamous Italian families is equal to 6.4%, while in mixed families with at least one Italian spouse it increases to 7.5%. Larger differences across families are uncovered for the remaining ethnic groups, i.e. the gap for the European group is of 3.7%, 3.9% for Other European, 7.3% for North Africa, 6.6% for other Africa, 5.4% for West Asia, 4.1% for East Asia and 3.5% for America. Moreover, by looking at ethnic group variability across homogamous unions, we report evidence that the separation rate of Italian families, as the majoritarian group, is higher with respect to those of homogamous minorities. The evidence is consistence with theoretical predictions, because the horizontal socialization makes dissolution choices less riskier for the majoritarian group in terms of children cultural transmission. We exclude that differences in instability across ethnic groups are driven by a carry-over of preferences towards marital dissolution of migrants, as the Italian divorce rate is particularly low compared to the average divorce rates in migrants countries of origin [16]. The evidence, discussed so far, documents the existence of strong preference for homogamy along cultural-ethnic lines and show a differential in household investments and stability across family types. The scope of the research is to investigate the role of the cultural socialization channel, in explaining the observed patterns.

We proxy the cultural-ethnic transmission by the language spoken at home by parents with children. By looking at the probability of speaking Italian at home, we show that socialization rates are quite low in Italy, in particular for homogamous unions of spouses belonging to minorities. Homogamous minorities families have a strong preference toward

[16] Divorce rates data come from the World Marriage Data 2008. United Nations, Department of Economic and Social Affairs, Population Division (2009).

the transmission of their own language to children, hence Italian frequencies are lower compared to those of heterogamous families. For instance, the probability that a European parent speaks Italian with his/her child is equal to 38.1% in a homogamous marriage, whereas it increases to 90.7% in a heterogamous marriage, and it remains quite high to 69.8% excluding mixed marriages with an Italian spouse. Across group variability is particularly high. The probability that a South-West Asia parent speaks Italian with his child is equal to 21.4% in a homogamous marriage, whereas it increases to 92.1% in a heterogamous marriage. Estimates point out that homogamous and heterogamous couples are endowed with different vertical technologies of cultural transmission. Conflicting incentives in the socialization of children for mixed unions lead to less efficient vertical transmission technologies as compared with homogamous families, which benefit from coordinated incentives.

Italian language socialization frequencies by spouses ethnic group and marital status are consistent with our model. Divorced homogamous families have less efficient socialization technologies as compared to married homogamous families. For example, the probability that a child speaks Italian is equal to 36.4% in a European homogamous married family and it increases to 43.6% for the same family under divorce. Vice versa, divorced heterogamous families have more efficient socialization technologies as compared their married counterpart, hence for a comparable European heterogamous families, the probability that a child speaks Italian is equal to 92.2% and 79.5%, respectively, in case of marital stability or not.

2.6 Econometric Framework

2.6.1 The Structural Model

Supported by the descriptive evidence, the analysis that follows will further examine the relationship between cultural transmission preferences and household choices. We specify the structural model in this section, introducing relevant assumptions and functional form parametrizations for the econometric implementation of the theoretical model. Afterwards, we describe our identification strategy and we introduce an appropriate estimation procedure. For instance, the structural model provides us reduced form of the theoretical moments as function of structural parameters, for given exogenous population distribution. We match theoretical moments implied by the model with their empirical counterparts observed in the data.

In the empirical application, the cultural-ethnic socialization is proxied by the language transmission. By looking at the language spoken at home by parents with children, we restrict the potential socialization probabilities to be positive, consistently with observed socialization frequencies. For instance, considering a marriage between a man i and a woman j, each child might be positively socialized to the trait i, j and I (Italian), with $i \neq I$ and

$j \neq I$, while remaining socialization probabilities are constrained to be zero. The restriction takes into account the fact that each child has a positive probability of speaking Italian at home with parents, independently of parents' ethnicities. Moreover, the same child has a positive probability of speaking the mother and/or the father foreign language. Thus, in a marriage of type $\{i,j\}$ in divorce status d in a population distribution q^i, residing in the province g, the expected utility from socialization of male m, per child is:

$$W^i_{m,g}(\tau;\{i,j\},d,q^i_g) = P_i V^i_i + P_j V^i_j + P_I V^i_I \quad \forall i \neq j, \quad \forall i,j \neq I$$

and respectively for the mother. Implicitly we assume that homogeneous Italian families are endowed with a perfect socialization technology, as verified by the data.

We model the role played by the society within the transmission process, relaxing the initial assumption of unbiased horizontal socialization frequencies. We explicitly introduce a positive bias toward the Italian socialization, in such a way that $Q^I_g > q^I_g$. Referring to the most general marital union with $i \neq j$ and $\forall\, i,j \neq I$, the horizontal matching probabilities are:

$$Q^i_g = \frac{q^i_g}{\sum_k q^k_g} \quad \forall i \neq I, \quad \forall k, \quad \forall g \tag{2.8}$$

$$Q^j_g = \frac{q^j_g}{\sum_k q^k_g} \quad \forall j \neq I, \quad \forall k, \quad \forall g \tag{2.9}$$

$$Q^I_g = 1 - Q^i_g - Q^j_g \quad \forall i \neq j, \quad \forall i,j \neq I, \quad \forall g \tag{2.10}$$

and they satisfy the accounting constraint: $\sum_i Q^i_g = 1\ \forall g$.

Moreover, we specify the socialization cost and the fertility cost functions as:

$$C(\tau_g) = \sigma_\tau \left\{ \lambda_\tau \frac{1}{2} (\tau_g)^2 + (1-\lambda_\tau) \left(\exp\left(\frac{\tau_g}{1-\tau_g} \right) - 1 \right) \right\} \tag{2.11}$$

$$C(n_g) = \sigma_n \left\{ \lambda_n (n_g)^\xi + (1-\lambda_n) \left(\exp(n^\xi_g) - 1 \right) \right\} \tag{2.12}$$

given $\xi > 1$. In this context, ξ mean to capture the dependence of fertility costs on childbearing decisions. The parametrizations of socialization and fertility cost functions guarantee that they are increasing and convex functions in the parents socialization efforts and childbearing choices, respectively, and they satisfy regularity Inada conditions for interior solutions. [17].

[17] For all $q^i \in (0,1)$:

$$\lim_{\tau_g \to 0} \frac{\partial C(\tau_g)}{\partial \tau_g} = 0, \quad \lim_{\tau_g \to 1} \frac{\partial C(\tau_g)}{\partial \tau_g} = \infty,$$

Finally we assume a standard logistic distribution for the quality of the marital union θ. Recalling the marital dissolution choice problem in its additive random utility form, we might define $\pi\left(\{i,j\},n,q^i\right)$ as:

$$\begin{aligned}\pi\left(\{i,j\},n,q^i\right) &= Pr[n\tilde{W}(\{i,j\},1,q^i) > n\tilde{W}(\{i,j\},0,q^i)+\theta] \\ &= Pr[n\tilde{W}(\{i,j\},1,q^i) - n\tilde{W}(\{i,j\},0,q^i) > \theta].\end{aligned}$$

Given distributional assumption for θ, we explicitly derive $\pi\left(\{i,j\},n,q^i\right)$:

$$\pi\left(\{i,j\},n,q^i\right) = \frac{\exp\bar{\theta}}{1+\exp\bar{\theta}} \tag{2.13}$$

where

$$\bar{\theta} \equiv (n\tilde{W}(\{i,j\},1,q^i) - n\tilde{W}(\{i,j\},0,q^i)).$$

A complete description fo the reduced-form equations of the structural model is reported in Appendix (B.3).

2.6.2 *Estimation Strategy and Identification*

Our interest lies in estimating the parameters of the marital value function. The main parameters of interest in the model are the cultural relative intolerance parameters: $\Delta V^{ij} \equiv V_i^i - V_j^i$ for all i,j. These parameters measure the relative distance in values of a parent of trait i in having a child sharing is own cultural trait i rather than acquiring a different trait j. In particular, they mean to capture the strength of parents cultural transmission desire relative to any other potential trait. Finally, the model includes socialization and fertility cost function parameters $\sigma_\tau, \lambda_\tau$ and σ_n, λ_n, respectively; preference for fertility, in terms of the dependence of fertility costs on childbearing decisions $\xi > 1$. Let denote the vector of parameters $\boldsymbol{\beta}$.

The structural model provides us reduced form of the theoretical moments ($\tilde{\boldsymbol{\Pi}}(\boldsymbol{\beta})$) as function of structural parameters, for given exogenous population distribution q_g^i for all ethnic groups i and for all provinces g. Those reduced forms represent a mapping from $\boldsymbol{\beta}$ into $\varphi_g(\{i,j\})$, $n_g(\{i,j\})$, $\pi_g(\{i,j\})$, for all i,j and for all g, and $P_{k,g}(\{i,j\})$ for all i,j and k and for all g. We estimate model parameters via a minimum distance estimator, matching the vector of theoretical moments implied by the model, ($\tilde{\boldsymbol{\Pi}}(\boldsymbol{\beta})$) for a specified choice of parameters β, with their empirical counterparts observed in the data ($\hat{\boldsymbol{\Pi}}$). The minimum distance estimator $\hat{\boldsymbol{\beta}}$ minimizes the criterion function $Q_N(\boldsymbol{\beta})$, such as:

$$\lim_{n_g\to 0}\frac{\partial C(n_g)}{\partial n_g} = 0, \quad \lim_{n_g\to\infty}\frac{\partial C(n_g)}{\partial n_g} = \infty.$$

$$\hat{\boldsymbol{\beta}} = arg\min_{\boldsymbol{\beta}}[\hat{\boldsymbol{\Pi}} - \tilde{\boldsymbol{\Pi}}(\boldsymbol{\beta})]^{\intercal}\boldsymbol{W}_N[\hat{\boldsymbol{\Pi}} - \tilde{\boldsymbol{\Pi}}(\boldsymbol{\beta})],$$

given $\boldsymbol{W}_N$ a $r \times r$ weighting matrix, for r the number of reduced form parameters and N sample size.

Theoretical moments are calculated as follow. For a given value of the parameters $\boldsymbol{\beta}$, together with the exogenous religious shares q_g^i, first order conditions of the optimization problem in (2.3) pin down the socialization efforts $\tau_g(\{i,j\}, d, q_g^i)$. Given optimal efforts, we can compute the value of the indirect socialization utilities $\tilde{W}_g(\{i,j\}, 1, q^i)$ and $\tilde{W}_g(\{i,j\}, 0, q^i)$ for all i, j by province g. Conditional on a set of $\tilde{W}_g(\{i,j\}, 1, q^i)$ and $\tilde{W}_g(\{i,j\}, 0, q^i)$, and observed fertility rates $n_g(\{i,j\})$, we pin down the rates of separation $\pi_g\left(\{i,j\}, n, q^i\right)$. Then, we compute the fertility moments, which in turn allows to calculate the gains to marriage.

We consider the following empirical moments: marital gains (derived from the marital matching function) $\hat{\varphi}_g(\{i,j\})$ for each $\{i,j\}$ match and for all provinces; fertility rates $\hat{n}_g(\{i,j\})$ for each $\{i,j\}$ match and for all provinces; separations rates $\hat{\pi}_g(\{i,j\})$ for each $\{i,j\}$ match and for all provinces and socialization frequencies $\hat{P}_{k,g}(\{i,j\})$ for all i, j and k and for all provinces g. In particular, we compute $\hat{\varphi}_g(\{ij\})$ thanks to identification equation of the marital matching function in (2.7). We derive $\hat{\mu}_g(\{i,j\})$ from the bivariate distribution of marriages, cumulate over the period 1995-2012, for each province and we derive $\hat{\mu}_g(\{i,0\})$ and $\hat{\mu}_g(\{0,j\})$ from the population vectors by ethnic group, sex and marital status of individual Italian Census data in 2001 and 2011. See Table (**??**) for the overall distribution and Table (**??**), respectively. We computer fertility rates as the average number of children per match, including zeros. We evaluate separation rates as the probability of observing a separation during the period of analysis per match. Finally, we construct socialization frequencies, by looking at the probability of speaking a specified language at home per match. See Appendix (B.2) for a deep description of empirical moments computation.

We structural marital matching model extends the Choo and Siow (2006b)'s model to a multi-market framework, exploiting geographical variability of empirical frequencies. We presume that each province corresponds to a marriage market, hence being the relevant geographical unit of reference for marital selection and intra-household decisions.

Finally, we assume that individual preferences and cost functions share common characteristics across different markets. Specifically, cultural relative intolerance parameters are ethnic-group specific and vary across potential matches, but they are constant across markets. In addition, cost functions are independent both across ethnic-groups and marriage markets. Identification hinges on above parameters' restrictions and it relies on exogenous variability in the ethnic composition of the population across markets.

2.7 Conclusions

In this paper, we investigate marital matching along the cultural-ethnic identities of the spouses. The aim of the paper is to investigate a specific novel channel to explain the differential in household investments and stability across family types, namely cultural socialization.

We proposed a structural model to study the process of family formation and intra-household decision making in a context where ethnic differences between spouses matter. In particular, we consider a marriage matching model along the ethnic dimension. Parents choose fertility, investments in the cultural socialization of children, and possibly divorce. We show that the socialization channel implies complementarity in cultural traits of spouses and we show that differences in transmission technologies have implications for fertility and divorce patterns, from a theoretical perspective.

The empirical investigation exploits variability in ethnic identity across migrants, within Italy. We use administrative individual level data on marriages, separations and births, from Italy and we construct a unique quasi-longitudinal dataset, merging marriage records with birth and separation records. The dataset covers the universe of marriages formed in Italy from 1995 to 2012. Socialization frequencies come from the Condition and Social Integration of Foreign Nationals Survey (2011-2012), exploiting information on language transmission. We propose to estimate model parameters (cultural intolerance, integration and socialization technology parameters) via a minimum distance estimator. In line with theoretical predictions, the observed intermarriage, fertility, separation and socialization rates are consistent with strong preferences of parents toward cultural socialization of children to their own ethnic identity, proxied by language transmission.

Chapter 3

The Price of Citizenship: The Effect of EU Enlargement on Marital Matching in Italy [1]

Abstract

This research proposes and estimates a transferable utility marital matching model along the cultural-ethnic identity of spouses. The analysis focuses on interethnic unions, between one Italian spouse and one non-Italian spouse and exploits administrative individual level data on marriages formed between 1995 and 2012 in Italy. I argue that gains to interethnic marriage depend on both cultural preferences and legal status or citizenship acquisition motives. I estimate the *price of citizenship*, taking advantage of the change in the legal status regulation, caused by the exogenous enlargement of the EU in 2004 and 2007 to include East European countries. Results show that, only after the EU enlargements, legal status acquisition induced a huge decrease in the gains to intermarriage for migrants of new EU member countries. Further, I provide evidence in favor of cross-ethnic marital substitutability. Finally, heterogeneous effects across marriage markets are suggestive that intermarriages driven by legal status motives trade-off economic-labor opportunities.

Keywords: Ethnic intermarriage, Marital matching, EU enlargement, Legal Status acquisition.

JEL Classification: J11, J12, J15.

[1] This chapter is part of a broad research investigation joint with Jérome Adda (Bocconi University) and Paolo Pinotti (Bocconi University). I am grateful to Lorena Viviano and Luigi Di Gennaro at the ADELE Laboratory in Milan for their support and suggestions.

3.1 Introduction

The regulation of migration inflows and of migrants' legal status and citizenship acquisition are crucial and debated questions in Western countries which have received a renewed attention following recent political results in the UK and the US. A large body of literature has investigated the effects of immigration and immigration regulation policies on various political and economic outcomes. The contribution of the present research is to analyse the consequences of migrants' legal status acquisition on their marital matching choices. The direction of this relationship is not only still an open question, but it is also a fundamental feature of the process of socio-economic integration of new minorities.

The research focuses on ethic intermarriages, i.e. marriages where spouses share different cultural-ethnic traits. I exploit variability in ethnic identity across migrants within Italy, in accordance with the continuing increase of migratory inflows. Indeed, Italy experiences different waves of immigration with heterogeneity both in time and ethnic composition. For instance, at the beginning of 2014, the number of foreign residents registered in Italy was more than 4.9 million and they accounted for 8.1% of the total resident population. This is in comparison to 2013, when the same percentage was around 7.4% and as little as ten years earlier, in 2003, it was 2.6%. Migratory inflows have affected the Italian marriage market. For instance, the percentage of interethnic marriages nearly tripled over the period from 1996 to 2013, especially for marriages between a native husband and a foreign-born wife. The increase accelerated mainly at the beginning of the period and smoothly converges over time, despite the sustained raise of the foreign population, as reported in Figure 3.1. This study empirically investigates the extent to which the flattening of the intermarriage rates is explained by the change in the legal status regulation driven by the recent enlargements of the European Union (EU) to East European countries.

I nest the empirical analysis inside a structural marital matching framework. The research proposes and estimates a marital matching model along cultural-ethnic lines. In particular, I advance the hypothesis that gains to intermarriage depend on both cultural attributes of spouses and legal status or citizenship acquisition motives. To overcome the endogeneity of legal status acquisition, I estimate the price of citizenship, taking advantage of the change in the EU citizenship regulation, driven by the exogenous EU enlargements to East European countries in 2004 and 2007. I document that legal status or citizenship acquisition affects marital matching choices, reducing intermarriage incentives.

The study of heterogamous marriages and its evolution over time has been a subject of much discussion amongst scholars across a variety of disciplines [2]. The present research builds on these premises and contributes to the literature analysing intermarriage

[2] Kalmijn et al. (2005); Fu and Heaton (2008); Fryer (2007) and Schwartz (2013).

formation choices within a structural matching framework. I propose and estimate a static transferable utility marriage matching model along the cultural-ethnic identity of spouses in a frictionless market. Preferences are described by an additive random utility model à la Choo and Siow (2006b). Marital gains are characterized both by a systematic component, which depends on cultural attributes of spouses, and idiosyncratic individual components. Within this framework, I advance the hypothesis that systematic gains to marriage are the result of different incentives. On the one hand, individuals have preferences over spouses' cultural attributes, rationalized by complementaries in the household production function. On the other hand, non-native individuals might have preferences toward mixed marriages, arising from legal status and/or citizenship acquisition motives. A stochastic structure is also introduced at the level of individual preferences [3]. In line with Choo and Siow (2006b)'s contribution and more recent extensions [Galichon and Salanié (2015) and Chiappori, Salanié, and Weiss (Chiappori et al.)], the model provides exact non-parametric identification of the gains to marriage along ethnic lines, for all potential matches.

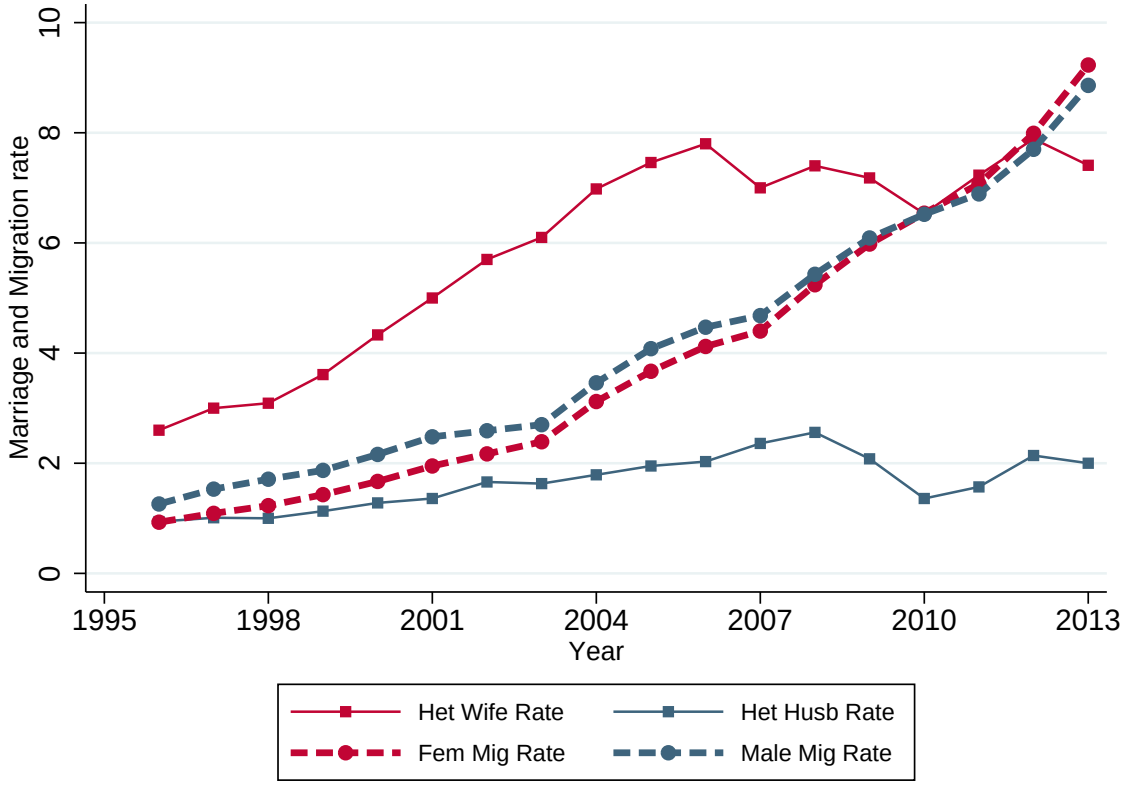
From the empirical point of view, I estimate the gains to marriage of interethnic unions formed in Italy, thanks to administrative data on marriages and individual Census data. Differences in cultural traits are evaluated with respect to the ethnic identity of spouses, proxied by the country of origin. I focus on interethnic unions between one Italian spouse and one non-Italian spouse and estimate the gains to intermarriage, exploiting both time variability and geographical variability across provinces. The identification of gains to intermarriage allows to investigate how changes in the migrants' legal status acquisition affect their marital matching decisions. In particular, I estimate the *price of EU citizenship*, taking advantage of the change in the EU citizenship regulation driven by the exogenous EU enlargements to East European countries in 2004 and 2007 [4]. Indeed, the EU enlargements modified the legal status of migrants in Italy coming from new EU member countries, potentially affecting their systematic marital preferences.

I estimate the effect of legal status acquisition, exploiting both time variation (comparing gains formed before and after the EU enlargements) and across countries variation (comparing new EU member countries versus non new EU member countries) via a difference-in-differences strategy. Namely, I evaluate the differential change in gains to intermarriage before and after the EU enlargements, comparing intermarriages of migrants from new EU member countries with migrants for whom the legal status regulation did not change with the EU enlargements. Results show that legal status acquisition induced a huge decrease in the gains to intermarriage for migrants of new EU member countries. The heterogene-

[3] Individual-specific idiosyncratic components aim to capture non-monetary returns from marriage and allows to rationalize heterogeneity in optimal sorting decisions of observationally equivalent individuals.

[4] In May 2004, Cyprus, Estonia, Latvia, Lithuania, Malta, Poland, the Czech Republic, Slovakia, Slovenia and Hungary joined the EU; Romania and Bulgaria joined in January 2007.

Figure 3.1: Evolution of Intermarriage and Migration Rates, Italy (1996-2013)



Source: Marriage records from vital statistics registries (1996-2013), Italy. The Figure reports the evolution over time of the intermarriage rate, as the ratio between the number of intermarriages formed in a given year over the total number of marriages formed in Italy in the same year. The intermarriage rate is reported separately for intermarriages between an Italian husband and a foreign-born wife (Heterogamous wife) and an Italian wife and a foreign-born husband (Heterogamous husband). *Source:* Movements of total foreign residents (1996-2013), Italy. The Figure reports as well the migration rate, as the ratio between the number of migrant residents over the total resident population in Italy.

ity in responses to the EU enlargements points to the important role that the legal status acquisition plays in the valuation of intermarriage by those minorities that are far behind the majority group, from both an economic and a socio-cultural perspective. The effect anticipates the introduction of the reform, and it discloses starting from the EU Council Decision. By looking at fully-flexible estimates, I do not observe any trends of the estimated interaction effects during the period immediately prior to the EU enlargements. Moreover, I provide evidence in favour of cross-ethnic marital substitutability. Exploiting different counterfactual groups, the analysis accounts for spillover effects generated in the market, in terms of cross-ethnic marital substitutability. Finally, I document the presence of heterogeneous effects of legal status acquisition across marriage markets. I document a larger decrease in gains to intermarriage in markets with higher economic-labor prospects for legalized migrants, suggesting that intermarriages driven by legal status motives trade-off

economic-labor opportunities.

A deep investigation of marital substitution patterns is a particularly important question. The project develops in this direction, proposing and estimating a formal multidimensional marital matching framework, along age and educational spouses characteristics.

The rest of the chapter is organized as follows. Section 3.2 provides a review of the literature. Section 3.3 describes the EU enlargements to East European countries. Section 3.4 outlines the theoretical framework of reference. Section 3.5 describes the data used in the empirical analysis, while Section 3.6 presents the empirical strategy and related results. Finally, Section 3.7 concludes and briefly points to future research directions.

3.2 Literature Review

Starting from Becker's seminal contributions [1973, 1974, 1991] [5], family economics research developed, investigating *who marries whom and why*. Several studies looked deeply into the assignment process of spouses, when competition for potential partners in the market affects individual choices [6]. For instance, the decision to enter a specific union depends on the value of the specific match, but also on the whole range of potential matches in the market.

In particular, some studies focused on sorting mechanisms along age. Choo and Siow (2006b) studied how the national legalization of abortion in the US affected the gains to marriage, especially for women of childbearing age. Similarly, Choo and Siow (2006a) and Choo (2015) investigated the effect of demographic changes such as the coming into marriage age of the baby boom generation or the gender differential in the mortality rate, affect the marriage market and the optimal matching process. Chiappori et al. (2012) estimated the marital trade off between income and body mass index, considering a bidimensional matching model. Chiappori et al. (2009) and Chiappori, Salanié, and Weiss (Chiappori et al.) analysed the educational assortativeness to examine whether and how spouses' asymmetries in returns to education within marriage (i.e. marital college-plus premium) might be responsible for gender differences in high-education investments.

Within these important contributions the cultural dimension has been greatly overlooked, despite the recognition of its relevance in explaining economic outcomes and the strong persistence of cultural traits across generations [f.e. Bisin and Verdier (2000); Bisin et al. (2004); Fernández et al. (2004); Doepke and Zilibotti (2008) and Fernández (2013)]. As an exception, Hitsch et al. (2010) analyse a multidimensional matching framework along

[5] Becker (1973); Becker (1974). See also Becker et al. (1977) and Becker (1991).

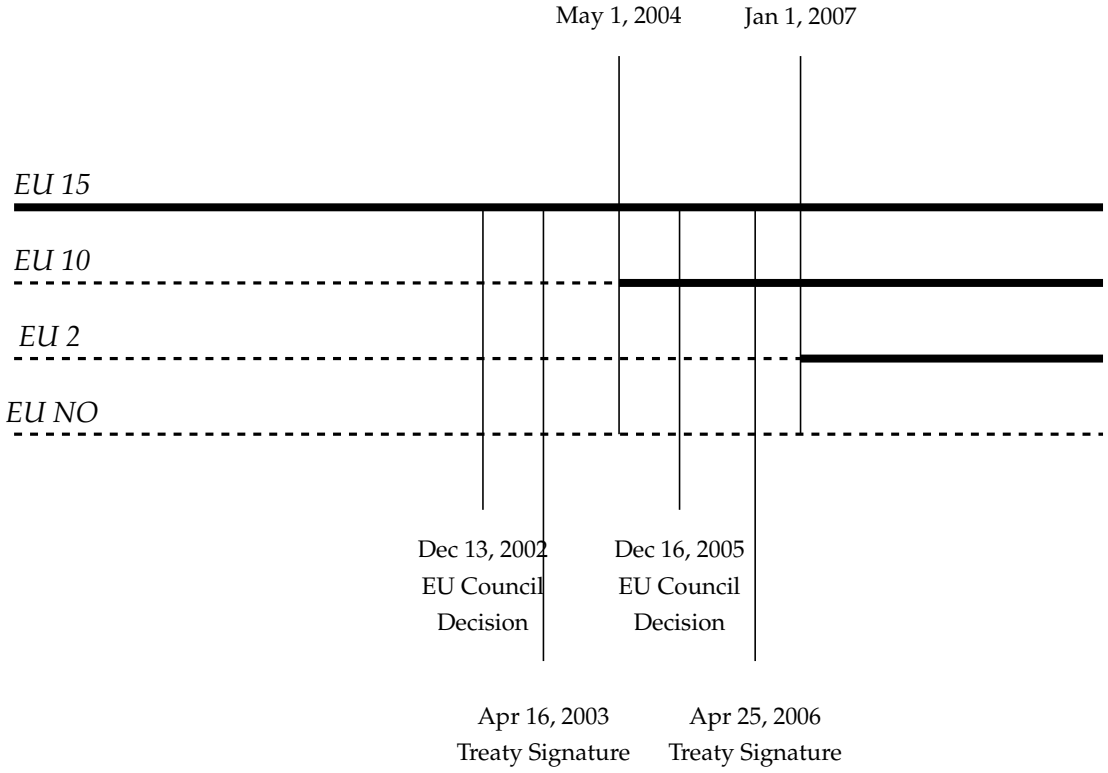
[6] For a survey of the literature see Browning et al. (2014) Chapter 7 and Chiappori and Salanie (2016).

age, education and race, exploiting data from online dating websites. The feature of online dating allows to investigate sorting patterns, discriminating between preferences and market frictions. They document a strong sorting on online matches along observable dimensions, which is attributed to preferences. Moreover, they provide evidence that online sorting qualitatively predicts sorting on actual marriage. Similarly, Wong (2003) proposes a model of match and searching along ethnic-racial lines where complementarity in marital output is reduced because interracial mating is considered taboo. Results point out that mating taboos explain the observed low interracial marriage rate compared to opportunities or endowment dimensions. This project relates and aims to contribute to this recent matching literature, bridging the current gap between the cultural and family economic literature.

By analysing the effect of the EU enlargements on the marriage market, this work also contributes to a large literature looking at the economic consequences of migrants' legal status acquisition. Despite several studies providing empirical evidence that legal status acquisition improves the labor market opportunities of migrants [Amuedo-Dorantes et al. (2007) and Lozano and Sorensen (2011)) and lowers crime rates [Mastrobuoni and Pinotti (2015) and Pinotti (2017)], less attention has been devoted to quantify its impact on socio-demographic dimensions. The study of Azzolini et al. (2015) contributes in this direction, estimating the effect of the EU enlargements to East European countries on marital choices through a reduced form approach. They exploit a synthetic control method to derive the counterfactual intermarriage rates for new EU member countries after the EU enlargements, thanks to observed intermarriage rates of not EU countries. The present research expands the scope of Azzolini et al. (2015) analysis, providing an identification of the effect of the legal status acquisition by exploiting a structural estimation of the gains to intermarriage. Moreover, I acknowledge the fact that the exogenous variation in legal status acquisition for new EU member countries might have lead to spillover effects. I provide preliminary evidence in favour of cross-ethnic substitution, taking into account the general equilibrium effects.

Finally, this research speaks to a large empirical literature on migrants assimilation. Indeed, interethnic marriages have been widely used by the academic literature on migration to study and evaluate the process of cultural assimilation and integration of immigrants into the host countries [Meng and Gregory (2005); Constant and Zimmermann (2008); Constant et al. (2009) and Algan et al. (2012)]. The present contribution underlines the importance of the legal status motive to describe intermarriage choices.

Figure 3.2: Timeline of the EU Enlargements to East European Countries



3.3 EU Enlargements to East European Countries

This investigation studies the effect of the legal status acquisition on marital matching choices. The causal identification of this relationship is hampered by the fact that the acquisition of the legal status and the naturalization process are endogenous, as they positively correlate with both observable and unobservable individual characteristics, which might affect as well matching choices. For instance, migrants who are culturally more integrated into the host society are more likely to acquire the legal status because of better labor opportunities, and at the same time they might be more prone to intermarry. I take advantage of the exogenous enlargements of the EU to East European countries in 2004 and 2007, to overcome the endogeneity issue. The enlargements provide exogenous variation in the legal status acquisition of migrants of new EU member countries.

Today, the EU is an economic and political partnership among 28 countries. However, this configuration is the result of a long process of subsequent enlargements, which is still *in fieri* as several countries are negotiating admission conditions [7]. Indeed, the EU was for-

[7] A complete timeline of the EU enlargements, from the initial signature of the Treaty of Rome in 1957, is reported in Figure C.1, while a detailed description of the overall process could be found here: https://ec.europa.eu/neighbourhood-enlargement

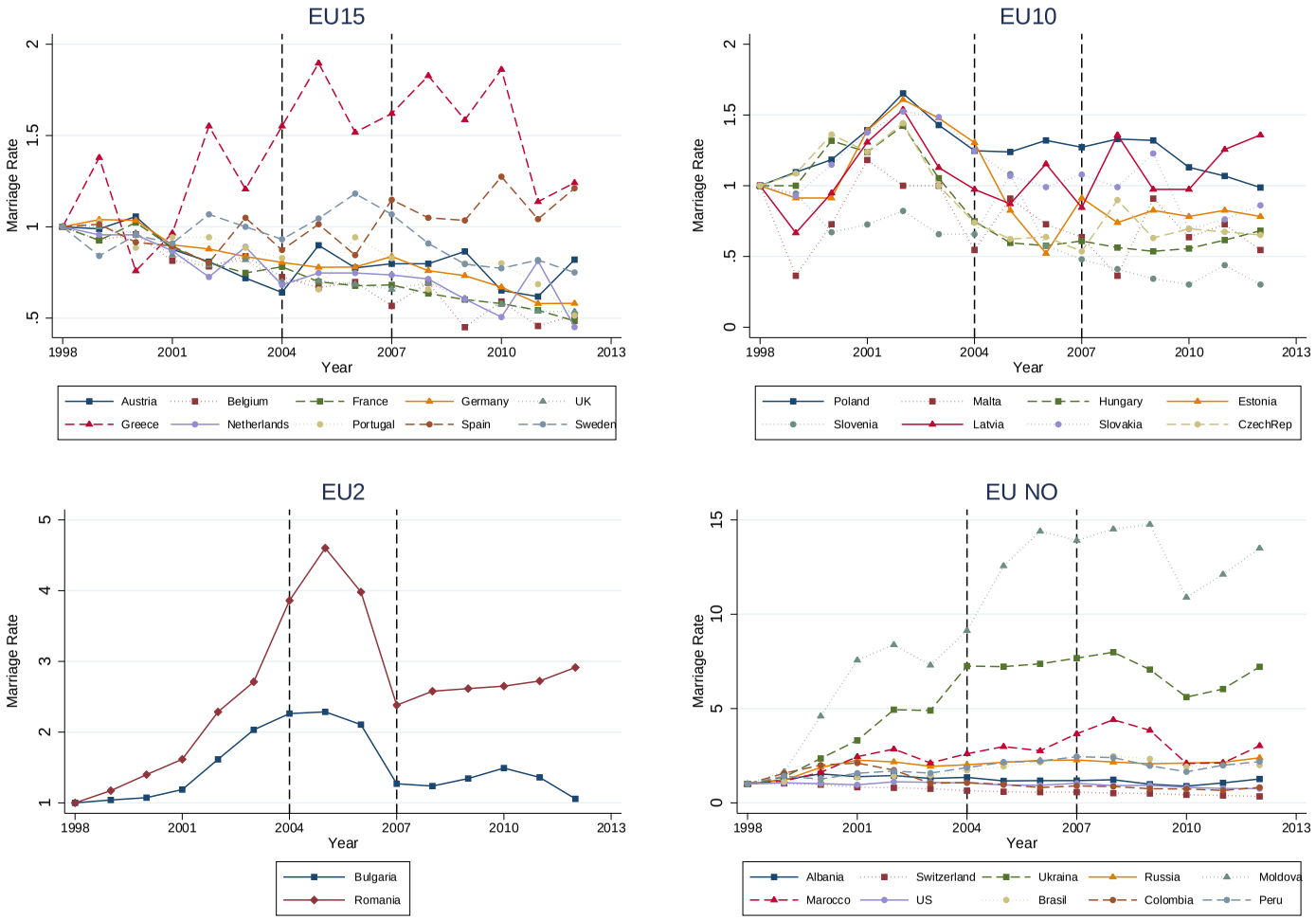
mally established when the Maastricht Treaty came into force on November 1, 1993, counting at the time 12 member countries: Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, and the United Kingdom. In 1995, Austria, Finland, and Sweden joined the EU. Then, the EU extended to East European countries in two rounds of enlargement. The first one, on May 1, 2004 represented the single largest EU enlargement in terms of people and number of countries, involving ten countries: Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovakia and Slovenia, and it . The second enlargement took place on January 1, 2007, when Bulgaria and Romania joined the EU. The latest enlargement to Croatia on July 1, 2013 leads to the current configuration. Let *EU*15 denote the former EU countries, *EU*10 denote the countries joining the EU in 2004, *EU*2 denote the countries joining EU in 2007, and *EU_NO* the countries that were not part of the EU after 2007. Figure 3.2 provides a graphical representation of the timing of the two enlargements. The timeline reports by chronological order: the date of the EU council agreement, the date of the Treaty signature and the date of accession for both enlargements. In both cases, the announcement of the EU Council decision largely anticipates the final accession. Estimates account for anticipatory effects, allowing for multiple cutoff definitions.

The enlargements removed the EU entrance restrictions faced by migrants of new EU member countries and conferred them the EU citizenship, i.e. the right to move and reside freely within the territory of the EU, as averse to previous temporary residence permit. As described in Mastrobuoni and Pinotti (2015), the admission to the EU allows migrants of new EU member countries to (i) look for a job in any other country within the EU; (ii) work there without needing any permit; (iii) live there for that purpose; (iv) stay until the end of the employment relationship, and (v) enjoy equal treatment as natives in access to employment, working conditions, and all other social and tax advantages that may help integration inside the host country.

For the purpose of this research, I focus only on intermarriages between one Italian spouse and one non-Italian spouse. Figures 3.3 and 3.4 plot the evolution of the intermarriage rate over time by political group, separately for *heterogamous wife* and *heterogamous husband* intermarriages[8]. The figures provide preliminary evidence in favour of a negative effect of the legal status acquisition on intermarriages. While *EU*15 countries show a constant declining trend in intermarriages, a sharp decrease in the number of intermarriages is observed for Romania and Bulgaria, starting from 2005. A similar negative decline is observed among *EU*10 countries, especially for Czech Republic, Hungary, Estonia and Latvia starting from 2003, even if less pronounced.

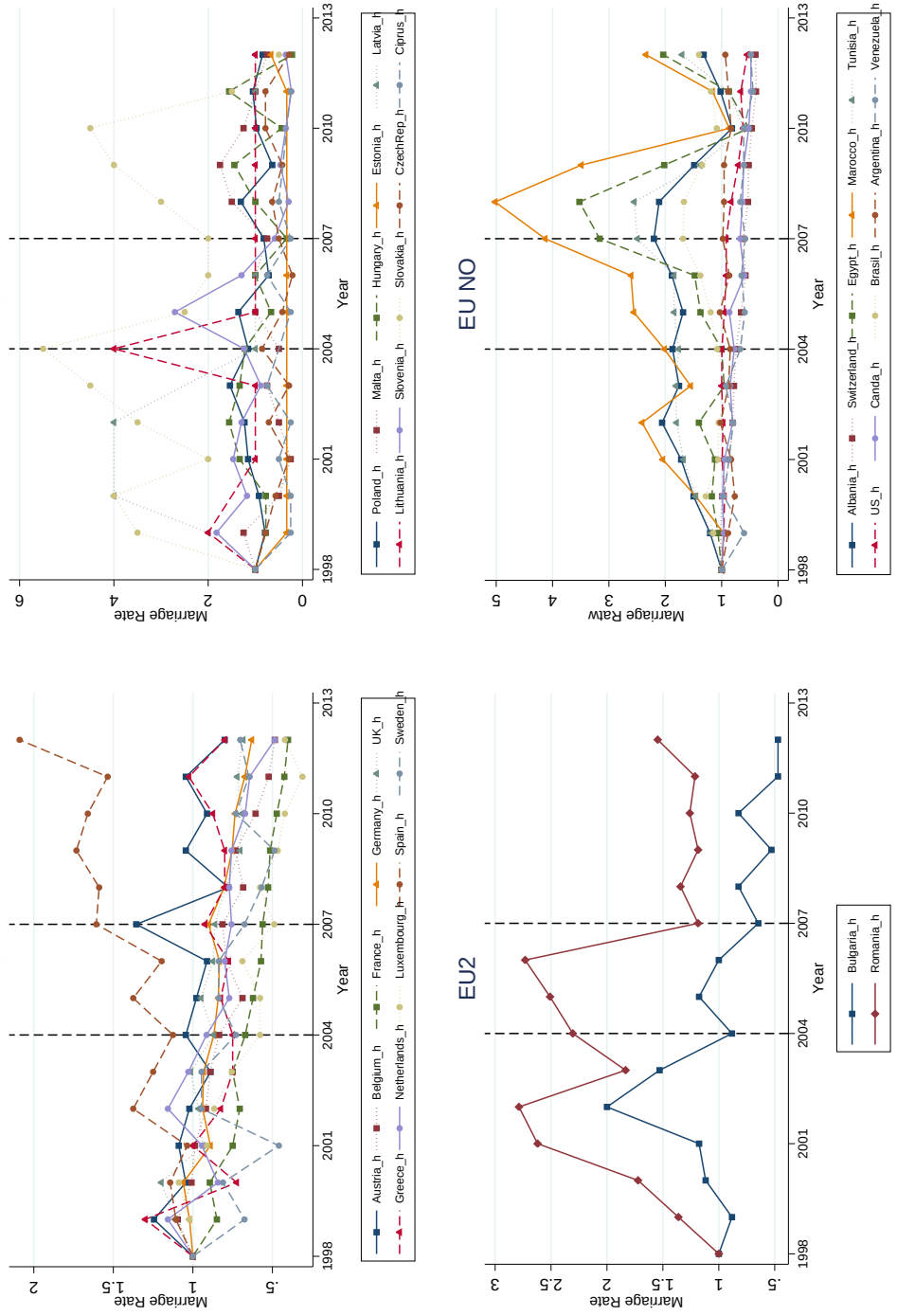
[8]The variable *heterogamous wife* denotes intermarriages between a foreign-born wife and a native husband, viceversa *heterogamous husband* intermarriages are unions between an Italian wife and a foreign-born husband. A description of the data is presented in Section (3.5).

Figure 3.3: Heterogamous Wife Intermarriages over Time, Italy (1998-2012)



Source: Marriage records from vital statistics registries (1998-2012), Italy. The Figure reports the evolution over time of the intermarriage rate for *heterogamous wife* unions, separately for *EU*15, *EU*10, *EU*2 and *EUNO* most relevant countries. the intermarriage rate is calculated as the ratio between the number of intermarriages formed in a given year over the number of marriages formed in 1998. The classification of countries is described in Section (3.3).

Figure 3.4: Heterogamous Husband Intermarriages over Time, Italy (1998-2012)



Source: Marriage records from vital statistics registries (1998-2012), Italy. The Figure reports the evolution over time of the intermarriage rate for *heterogamous husband* unions, separately for *EU*15, *EU*10, *EU*2 and *EUNO* most relevant countries. the intermarriage rate is calculated as the ratio between the number of intermarriages formed in a given year over the number of marriages formed in 1998. The classification of countries is described in Section (3.3).

Conversely, I observe an increase in intermarriages with women from Moldova, Ukraine and Morocco. Despite the fact that many EU countries preserved temporary EU entrance restrictions to the movement of migrant workers, the enlargements push migratory inflows. The increase in immigration might mechanically translate into an higher number of intermarriages. Hence, I study the intermarriage response to the EU enlargements within a marital matching framework, which delivers a structural identification of the gains to intermarriage, taking into account variations in the population distribution as well as in preferences.

3.4 Theoretical Framework: Marital Matching Model

3.4.1 The Setup

I consider a transferable utility (TU) marital matching model along the cultural-ethnic identity of spouses in a frictionless marriage market. The TU assumption captures the idea that spouses implicitly transfer utility between each other in absence of transition costs. Transfers are endogenously determined as equilibrium outcomes, as they depend not only on the quality of the specific match, but also on the whole set of available opportunities in the market. Moreover, in absence of frictions in the market, I assume that any individual has complete and costless information about all subjects present in the market and their observable characteristics. I study a one-to-one matching framework in a two-sided large marriage market, with a population of M men and F women. Men and women are heterogeneous in their cultural-ethnic identity. The model accounts for I types of men $m \in M$ and J types of women $f \in F$ [9]. Given men and women heterogeneity in cultural traits, $I \times J$ marriages are potentially observed in the market. Let N denote native individuals (i.e. Italians).

A *matching* defines who is matched with whom and who remains unmatched. Specifically, a matching is a measure μ_{ij} over the $I \times J$ space, denoting the probability of observing a match between a man of type i and a women j from the reference population. Similarly let μ_{i0} and μ_{0j} denote the probability density of unmatched men i and women j in the same population, respectively. The optimal matching must satisfy the following feasibility and non-negativity constraints:

[9] I consider ethnicity as an inherited trait, i.e. its formation is defined by intergenerational transmission differently from acquired traits, such as schooling. Both ethnicity and education influence marital selection, but while acquired traits are affected by marriage perspective inherited traits are independent from these feedback effects [Chiappori et al. (2009)].

$$\begin{aligned} &\sum_j \mu_{ij} + \mu_{i0} = m_i \quad \forall i \in I, \\ &\sum_i \mu_{ij} + \mu_{0j} = f_j \quad \forall j \in J, \\ &\mu_{ij} \geq 0, \mu_{i0} \geq 0, \mu_{0j} \geq 0. \end{aligned} \tag{3.1}$$

3.4.2 Systematic Preferences

Preferences are described by an additive random utility model à la Choo and Siow (2006b), where marital gains include both a systematic component and idiosyncratic individual components. I will start presenting the marital systematic components, while a detailed discussion of the idiosyncratic individual tastes follows in Section 3.4.4. In particular, I argue that systematic gains to intermarriage are the result of different incentives. On the one hand, individuals have preferences over the cultural attributes of their potential spouses. Let α_{ij} (γ_{ij}), be the surplus that a man i (woman j) receives in marrying a woman of type j (a man of type i). This systematic return is reduced (increased) by equilibrium transfers τ_{ij}[10]. Notice that both dimensions are common for all i, j potential matches. On the other hand, non-native individuals might have preferences toward intermarriages with natives, arising from legal status and citizenship acquisition motives. Let the value of the legal status acquisition for a foreign-born man $i \neq N$ be denoted by β_{iN}, in case he marries a native woman N, and viceversa let β_{Nj} be the values of the legal status acquisition for women of type j in marring a native man N. Hence, the marital utility of man i (woman j) from matching with woman j (man i) is equal to [11]:

$$\begin{aligned} &\alpha_{ij} + \tau_{ij} + \beta_{iN} \mathbb{1}\{i \neq N, j = N\}, \\ &\gamma_{ij} - \tau_{ij} + \beta_{Nj} \mathbb{1}\{i = N, j \neq N\}. \end{aligned}$$

3.4.3 Optimal Stable Matching

The quantities: $\alpha_{ij}, \gamma_{ij}, \beta_{iN}, \beta_{Nj}, m_i, f_j$ are the primitives of the matching model. The endogenous equilibrium quantities, instead, are the matching patterns $\mu_{ij}, \mu_{i0}, \mu_{0j}$, as well as the equilibrium transfers τ_{ij}. Finally, let ϕ_{ij} denote the marital surplus generated assigning man i to woman j, which is shared endogenously between spouses. Thus, under the assignment μ_{ij}, the total surplus generated in the market is: $\sum_{ij} \mu_{ij} \phi_{ij}$. The *optimal matching* is the

[10] Transfers are not constrained to be non-negative, a priori.

[11] In the maximization problem, I normalize the outside option of remaining single to zero. I denote the choice set of man i by $J_0 = J \cup \{0\}$ and the choice set of women j by $I_0 = I \cup \{0\}$.

solution of the following welfare maximization problem over all potential matches, subject to feasibility and non-negativity constraints of equation (3.1):

$$\max_{\mu_{ij} \geq 0} \sum_{ij} \mu_{ij} \phi_{ij}. \tag{3.2}$$

Shapley and Shubik (1971) showed that the solution of the primal problem in (3.2) corresponds to a stable matching. A matching is stable if nobody would prefer to deviate from the assignment, i.e. neither a man i nor a woman j who are currently married would rather be single, nor a woman j and a man i who are not currently married together would both rather be married together than remain in their current situation (i.e. absence of blocking pairs, Chiappori and Salanie (2016)). The stable optimal assignment satisfies the following system of matching equilibrium conditions, where each man chooses the woman who maximizes his utility and at the same time each woman chooses the man who maximizes her utility. The maximization problems are, in turn:

$$\begin{aligned} u_i &= \max_{j \in J_0} \{\phi_{ij} - v_j\}, \\ v_j &= \max_{i \in I_0} \{\phi_{ij} - u_i\}. \end{aligned} \tag{3.3}$$

3.4.4 *Idiosyncratic Individual Tastes*

Compared to the deterministic specification of the model presented so far, I explicitly introduce a stochastic structure at the level of individual preferences, following Choo and Siow (2006b)'s seminal contribution and by further extensions [Chiappori and Salanie (2016), Chiappori, Salanié, and Weiss (Chiappori et al.), Galichon and Salanié (2015)]. Individual stochastic components are crucial to rationalize the heterogeneity in optimal sorting decisions of observationally equivalent individuals, otherwise unexplained [12]. In particular, I allow for individual-specific idiosyncratic tastes, reflecting heterogeneity in preferences across individuals [13]. Let ϵ_{ij}^{m} denote the idiosyncratic taste of man m with type i for a woman of type j. Respectively, let η_{ij}^{f} be the idiosyncratic taste of woman f of type j for a man of type i. They are assumed to be observed by individuals in the market, but unob-

[12] A different approach to rationalize the variability in optimal marital sorting, is to introduce some frictions in the market. The presence of frictions implies that any individual has imperfect and costly information about potential mates in the market, and it requires to define explicitly a meeting technology. Randomness on the meetings process guarantees that similar agents, in equilibrium, will have different types of partners. Following the seminal contribution of Shimer and Smith (2000), several authors have started to combine the search and the matching frameworks.

[13] The idiosyncratic components might also be interpreted as idiosyncratic non-monetary returns from marriage.

served from the econometric point of view. The introduction of this stochastic dimension translates the optimal matching problem in (3.3) into:

$$\begin{aligned} u_i &= \max_{j \in J_0} \{\phi_{ij} + \epsilon^m_{ij} + \eta^f_{ij} - v_j\}, \\ v_j &= \max_{i \in I_0} \{\phi_{ij} + \epsilon^m_{ij} + \eta^f_{ij} - u_i\}. \end{aligned} \tag{3.4}$$

The characterization of idiosyncratic preferences follows Choo and Siow (2006b)'s contribution in three main assumptions. First, I assume that idiosyncratic tastes are additive in preferences, letting the marital surplus generated within the i, j match be: $\phi_{ij} + \epsilon^m_{ij} + \eta^f_{ij}$. Secondly, I assume that idiosyncratic tastes satisfy a separability assumption, i.e. they only depend on observed characteristics of the potential spouses, excluding any potential complementarity between unobserved dimensions from both sides of the market. Finally, ϵ^m_{ij} and η^f_{ij} are assumed to be two independent and identically distributed random variables and to follow an extreme value type I distribution. Chiappori and Salanie (2016) and Galichon and Salanié (2015) show that the separability assumption of unobserved heterogeneity components allows to reduce the two-sided matching problem of equation (3.4) into a series of one-sided discrete choice problems, linked by the adding-up formula [14]:

$$\begin{aligned} u_i &= \max_{j \in J_0} \{\alpha_{ij} + \tau_{ij} + \beta_{iN} \mathbb{1}\{i \neq N, j = N\} + \epsilon^m_{ij}\}, \\ v_j &= \max_{i \in I_0} \{\gamma_{ij} - \tau_{ij} + \beta_{Nj} \mathbb{1}\{i = N, j \neq N\} + \eta^f_{ij}\}, \\ \text{st.} \quad \phi_{ij} &= \alpha_{ij} + \beta_{iN} \mathbb{1}\{i \neq N, j = N\} + \gamma_{ij} + \beta_{Nj} \mathbb{1}\{i = N, j \neq N\}. \end{aligned} \tag{3.5}$$

Given the specific distributional assumptions, it is possible to achieve exact non-parametric identification of the gains to marriage along ethnic lines, for all potential matches [15]. The

[14] The final condition in equation (3.5) might be interpreted as a complementary slackness condition associated with the Lagrange multiplier μ_{ij}. By complementary slackness, when the assignment probability is positive $\mu_{ij} > 0$, the sum of spouses indirect utilities correspond to the actual breakdown of the marital surplus; while $\mu_{ij} = 0$ when the constraint is slack. Similarly, the constraint represents the equilibrium stability condition. Namely, whenever the constraint hold with an equality, it implies that the value generated from the marital union in equivalent to the sum of spouses utilities, hence a matched is formed. On the opposite case when the constraint is not binding it implies that individuals might be better off under a different assignment, which contradicts the stability condition (because of the presence of blocking-pairs)

[15] Exact identification is achieved by the fact that the number of marital preference parameters to be identified is equal to the number of observed matches. In principle, the identification problem in matching models is particularly cumbersome, as recalled in Choo and Siow (2006b). In this context, with I types of men and J types of women, the number of preference parameters that characterize the system is equal to $2 \times I \times J$, i.e. in choosing to enter into a specific marriage, each man i has a taste for each type j women, and viceversa for women. On the contrary, we observe in the data only $I \times I$ matches plus the distribution of unmatched by gender: $I + J$. The TU model allows identification because halve the number of preference parameters, i.e. each match is described by marital output preferences: $I \times J$ instead of single individual preferences $2 \times I \times J$.

matching problem translates into a multinomial logit model and the joint marital surplus is identified from matching patterns:

$$\phi_{ij} = \ln \frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}}. \tag{3.6}$$

The econometric problem solves with the identification of systematic utilities, observing the bivariate marital distribution μ_{ij} and the vectors of unmatched individuals μ_{i0} and μ_{0j}. For instance, for each potential match i, j, the systematic gain to marriage is identified by the log ratio of the number of observed i, j marriages to the geometric average of those types (i,j) who are available, in line with Choo and Siow (2006b) [16]. In this sense, the marital matching function assigns larger gains to more frequently observed matches in the market, which are conveniently scaled by the geometric average of the number of unmatched individuals of the corresponding types, in order to account for the potential heterogeneity of the population distribution across types and gender.

3.4.5 *Comparative Statics on Legal Status Acquisition*

Recalling the discrete choice formulation of the matching problem in (3.5), it is easy to see how changes in preferences lead to variation in optimal matching patterns, and generate relevant general equilibrium effects.

Proposition 1. For $\beta_{iN} \geq 0$ $\beta_{Nj} \geq 0$, u_i and v_j are not decreasing function in β_{iN} and β_{Nj}, respectively.

Proof. Consider the expected maximization problem of man m with trait i, before he knows his realization of ϵ_{ij}^m. u_i is an increasing and convex function, corresponding to the expectation of the maximum of linear functions. By the envelope theorem, the effect on u_i of a marginal increase in the systematic component is equivalent to the conditional probability $\mu_{j|i}$, i.e. the probability for man m of choosing j conditional on being of type i. Given:

$$u_i = E\left[\max_{j \in J_0}\{\alpha_{ij} + \tau_{ij} + \beta_{iN}\mathbf{1}\{i \neq N, j = N\} + \epsilon_{ij}^m\}\right]$$

$$\frac{\partial u_i}{\partial \beta_{IN}} = Pr\left[\begin{array}{l} \alpha_{ij} + \tau_{ij} + \beta_{iN}\mathbf{1}\{i \neq N, j = N\} + \epsilon_{ij}^m \geq \\ \quad \alpha_{ik} + \tau_{ik} + \beta_{ik}\mathbf{1}\{i \neq N, k = N\} + \epsilon_{ik}^m, \quad \forall k \neq j \in J_0 \end{array}\right],$$

[16] Galichon and Salanié (2015) reach the same result and provide a generalization of the Choo and Siow (2006b)'s model applying convex analysis results to discrete-choice models. Maintaining the separability and large market assumptions, they provide an equilibrium characterization of the matching patterns as function of model primitives and derive identification of the joint surplus, in a more general distributional setting.

where

$$\mu_{j|i} = Pr\left[\begin{array}{l} \alpha_{ij} + \tau_{ij} + \beta_{iN}\mathbf{1}\{i \neq N, j = N\} + \epsilon^m_{ij} \geq \\ \quad \alpha_{ik} + \tau_{ik} + \beta_{ik}\mathbf{1}\{i \neq N, k = N\} + \epsilon^m_{ik}, \quad \forall k \neq j \in J_0 \end{array}\right].$$

Under the assumption that the legal status acquisition increases the value of an interethnic marriage, i.e. $\beta_{iN} \geq 0$ $\beta_{Nj} \geq 0$, granting the legal status to new EU member migrants independently from marital reasons, the EU enlargements induced a negative variation in gains to intermarriage. Thus, I expect to observe a reduction in the number of intermarriages with migrants from new EU member countries, in response to the exogenous change. Moreover, I expect the response to the legal status acquisition to be highly heterogeneous across migrants' ethnicities, with the magnitude of the negative effect to be inversely proportional to the socio-economic conditions of immigrants before the intervention.

3.5 The Data

I estimate gains to marriage along cultural lines from equation (3.6), using administrative Italian data on marriages (1997-2012) and individual Census data on population vectors (1991, 2001, 2011). I describe below the data used in the empirical analysis.

Marriages (1998-2012). I recover matching patterns by spouses' ethnic identity ($\mu_{ij} \quad \forall i \in I, \forall j \in J$), exploiting marriage records from municipality vital statistics registries. Vital statistics registries contain the universe of marriages celebrated each year in Italy from 1997 to 2012. Marriage records provide information about the date of marriage, the celebration ceremony (religious or civil), the municipality of celebration and the choice of the property regime by the spouses (community or separation property). The information provided for each spouse include: the date of birth, the municipality of birth, the municipality of residence at the time of marriage, the province of future residence of the spouses, the previous marital status, the education level, the employment status, and for migrant individuals the nationality and the country of origin [17].

Individual Census Data (1991, 2001, 2011). Population vectors of unmatched individuals by gender and by ethnic identity (μ_{i0}, μ_{0j}) come from individual Italian Census data for the years 1991, 2001, 2011. I select only adult individuals, of more than 18 years of age. Census data classify the marital status of an individual as: never married, at present married, separated *de facto*, legally separated, divorced or widowed. I consider an individual available

[17] In 1997 only the nationality of foreign-born individuals is provided, because the information of the country of birth was not collected. Despite the naturalization process was not frequent at the time, the nationality might underestimate the number of migrant individuals. To overcome the measurement error issue, I report estimates exploiting both a two-year marriage distribution for the period from 1997 to 2012, and a one-year distribution for the period from 1998 to 2012.

in case she/he is never married, legally separated, divorced or widowed.

For the sake of the empirical application, two features require to be highlighted. First, differences in cultural traits are evaluated with respect to the ethnic identity of spouses, proxied by the country of origin. I restrict the set of spouses' ethnicities to a finite number of political-ethnic groups, to have thicker cells of the marital distribution. The classification of countries reflects both the cultural-ethnic proximity of each country with respect to Italy, and its economic and socio-political proximity. The list of countries belonging to each political-ethnic group is reported in Table C.1. In particular, I divide the former EU member countries between Northern European countries, $EU15German$, and Southern European countries, $EU15Latin$. Countries that became part of the EU in 2004 are denoted by $EU10$, while those that entered the EU in 2007 are denoted by $EU2$. Following Mastrobuoni and Pinotti (2015), among the remaining East European countries, I differentiate between those countries that were negotiating admission conditions with the EU at the time of my analysis, $EUNext$, and all the remaining ones, $EUOther$. Finally, I split all other countries into North Africa, $AfricaN$; South-West and East Africa, $AfricaWES$; South-West Asia, $AsiaWS$; East Asia and Oceania islands, $AsiaE$; and the rest of developed OECD countries, $OECD$. In the rest of the empirical discussion, I focus my attention on the analysis of *heterogamous wife* intermarriages, which largely dominate the overall number of intermarriages formed in Italy [18]. Secondly, compared to the theoretical model, I enlarge the type space of individuals, considering a multi-market framework. I exploit both time variability and geographical variability across Italian provinces. I investigate both a one-year and a two-year marital distribution. Estimates from a two-year distribution, are reported as robustness, because it returns thicker cells. Moreover, I consider multiple local marriage markets at the province level. Let $t \in T$ denote the time dimension, while $p \in P$ refers to provinces. For instance, each match is now characterized by the ethnic identities of the spouses i and j, by the time-year of the marital union t, and the province of residence of spouses p at the moment of the marriage. Thus, gains to marriage $\phi_{ij,tp}$ belong to the space $I \times J \times T \times P$. I estimate gains to marriage along ethnic lines, from equation (3.6), for each i, j match, for each time-year t and for each province p, as follow:

$$\hat{\phi}_{ij,tp} = \ln \frac{\mu_{ij,tp}}{\sqrt{\mu_{i0,tp}\mu_{0j,tp}}}. \tag{3.7}$$

To estimate marital gains, I approximate the evolution of population vectors over time as follows. First, I calculate the difference in the number of unmatched individuals by gender and by ethnicity from 1991 to 2001 and from 2001 to 2011. Similarly, I consider the increase in migration patterns by ethnicity and I calculate the difference in the number of

[18] I leave the analysis of *heterogamous husband* intermarriages to future investigation.

foreign individuals from 1995 to 2010 and I compute for each year the corresponding share. Finally, I impute to each year a share of the overall difference of unmatched individuals, that is proportional to the migration variation share. The imputation allows to take into account the fact that different ethnicities experience heterogeneous rates of migration over time.

I recover the evolution of the number of marriages formed in Italy over time, as well as the number of available foreign individuals and the estimated gains to intermarriage, for each political-ethnic group. The time variation in the total number of marriages across the political-ethnic groups differs in magnitude. Intermarriages between Italian men and foreign women from new EU member countries highly decrease, whereas intermarriages between Italian men and foreign women from other countries decrease, but at lower rate. Conversely, the Figure evidences a disproportionate increase of available foreign women from new EU member countries, especially for $EU2$ women, as opposed to women belonging to the other groups. It is this differential in pre-post EU enlargements variation of marriage patterns across political-ethnic groups and the pre-post variation in population vectors that provide identification of the fall in gains to intermarriage, in response to the legal status acquisition.

3.6 Empirical Strategy and Main Results

3.6.1 Empirical Strategy: Difference in Differences

To estimate the effect of legal status acquisition on gains to intermarriage, I evaluate the relative change in gains to intermarriage before and after the EU enlargements. Unconditional estimates of the before-after variation in gains to intermarriage are reported in Table (**??**). Column (1) reports the average gains to intermarriage by ethnic-group. The lowest gains refer to intermarriages with North African women, followed by other African nationalities and South-West Asia nationalities. Columns (2) and (3) report the average gains to intermarriage before and after the first EU enlargement to East European countries in 2004, in turn, while the estimated before-after variation is reported in Column (4). Preliminary evidence shows a larger significant decrease in gains to intermarriage after the EU enlargements for $EU10$ and $EU2$ countries compared to neither countries reformed. Despite a large heterogeneity in the effects, results document a general decrease in gains to intermarriage for all ethnic-group considered, suggesting the presence of a systematic downward trend in intermarriage propensity.

The estimated difference in gains to intermarriage between new and not new EU member countries might overestimate the causal effect of the legal status acquisition, without taking into account the presence of systematic trends in intermarriage propensity. Par-

tialling out the time trend effect including a full set of year fixed effects, estimates confirm the results described. In addition, I proxy these systematic time-trends with the evolution in the gains to intermarriage observed in not reformed countries as counterfactual. By exploiting a difference-in-differences (DID) estimation strategy, I compare the relative change in gains to intermarriage in the post-enlargement periods relative to the pre-enlargements period between new EU member countries and countries for whom the legal status regulation did not change with the EU enlargements. Because of potential pre-announcements effects, I define the pre-post period, considering both the actual date of the accession of new EU member countries and the previous date of EU Council decision, as detailed in Figure (3.2). I estimate the following DID model:

$$\hat{\phi}_{ij,tp} = \alpha + \beta NewEU_j \times PostT_t + \lambda_j + \theta_t + \rho_p + \varepsilon_{j,tp} \quad i = N, \tag{3.8}$$

where $\hat{\phi}_{ij,tp}$ represents the estimated gains to intermarriage for the i,j couple (with $i = N$), at time t in the marriage market p. The variable $NewEU_j$ is a dummy equal to one for new EU member countries j and zero otherwise. I consider the two groups of new EU countries, $EU10$ and $EU2$, separately in the estimations. The dummy variable $PostT_t$ is equal to one for the years after the EU enlargements and zero otherwise, with T equal to 2004 and 2006 for $EU10$ and $EU2$, respectively. The variable $NewEU_j \times PostT_t$ represents their interaction. The estimated coefficient of interest β captures the difference-in-differences between the gains to intermarriage for new EU member countries as opposed to all other countries, before and after the EU enlargements, controlling for ethnic-specific, time and province heterogeneity. Ethnicity fixed effects, λ_j, account for systematic differences in gains to intermarriage across migrants' country of origin, related for example to systematic time-invariant differences in cultural values. Time fixed effects, θ_t, allow to control for systematic time trends in gains to intermarriage. Finally, I include province fixed effects, ρ_p, to account for time invariant unobserved heterogeneity in marital selection across marriage markets. Finally, $\varepsilon_{j,tp}$ is an error term, possibly serially correlated at the province level. I interpret β as the estimated effect of the legal status acquisition on gains to intermarriage. For instance, the coefficient measures the average additional change in gains to intermarriage after the EU enlargements relative to the average additional change before EU enlargement, experienced by countries that were affected by the legal status modification compared to those who were not. Table 3.1 reports the β coefficients for the $EU10$ and $EU2$ groups, at the baseline and considering potential anticipatory effects.

Results show that legal status acquisition induced a sizeable decrease in the gains to intermarriage of 54.4% for $EU10$ countries, and as large as 123.5% for $EU2$ countries. The heterogeneity in the responses to EU enlargements points to the conclusion that legal status acquisition largely affect the value of intermarriages for those minorities that are far

from the majoritarian group, both at the economic and socio-cultural level. By accounting for anticipatory effects of the EU enlargements, results suggest a decrease in the gains to intermarriage of 52.8% for $EU10$, and equal to 67.2% for $EU2$ countries. Interestingly, while for $EU10$ countries the anticipatory effects explain the entirety of the effect induced by the EU accession, for $EU2$ countries the anticipatory effects explain only half of the overall response to the legal status acquisition, leaving room for potential substitution effects having occurred in the meantime. While implicitly the model in (3.8) gives equal weight to each province independently of the marriage market size, similar results are obtained using number of marriages by province-weighted least squares estimates. In the rest of the chapter, I present WLS estimates, unweighed ones are reported in Appendix (C.1). Further, the results are robust to the exclusion of each political-ethnic group, ruling out the possibility that estimates are driven by a particular counterfactual group of countries.

I acknowledge the fact that the exogenous variation in legal status acquisition of $EU10$ and $EU2$ countries might have led to spillovers, in terms of cross-ethnic substitution effects. Hence, despite all countries having been indirectly affected by the enlargements, the intensity of the substitution might be different across groups, in relation to their economic and socio-cultural proximity with respect to new EU member countries. Despite strong evidence on substitution patterns is hard to provide without estimating the matching equilibrium, I shed some light in this direction estimating the model in (3.8) taking as counterfactual each political-ethnic group. As evidenced in Figure (6), point estimates are always negative and statistically significant, and they uncover a mild variability across political-ethnic groups, except for $NorthAfrica$. Panel (b) of Figure (6) shows qualitatively similar results, after accounting for anticipatory effects.

3.6.2 *Flexible Difference in Differences Estimates*

Equation (3.8) examines the average effect of legal status acquisition for new EU member countries, in response to the EU enlargements. The main identifying assumption is the temporal stability of potential outcomes, i.e. the conditional independence of the change in gains to intermarriage in absence of the EU enlargements. Had the EU enlargements not taken place, new EU member countries would have had the same trend in their gains to intermarriage as countries not affected by the enlargements. While this assumption is not directly testable, I provide evidence in favour of conditional independence, estimating a fully flexible DID model, as follows:

$$\hat{\phi}_{ij,tp} = \rho + \beta_t NewEU_j \times T_t + \lambda_j + \theta_t + \rho_p + \varepsilon_{j,tp} \quad i = N. \tag{3.9}$$

Different from equation (3.8), in this case I interact the dummy for $NewEU_j$ member countries with each of the year fixed effects (T_t is a dummy equal to one for each year t and

zero otherwise). The flexible DID model estimates a vector of year-specific β_t coefficients, which uncover the relationship between the legal status acquisition and the outcomes variable in each year, with respect to a baseline year, which I take to be 1998. According to the parallel trend assumption, I would expect the estimated β_t's to be constant for the years before the EU enlargements. Moreover, I expect the coefficients to decrease in magnitude after the EU enlargements, with a significant drop immediately after the EU enlargements. The patterns in the intermarriage response to the legal status acquisition can be clearly seen by plotting the β_t coefficients over time, as in Figure 3.7, with relative 95% confidence intervals. A clear pattern emerges from these estimates, and three important facts are worth to be mentioned. First, I do not observe any trends of the estimated interaction effects during the years immediately prior to the EU enlargements. Any systematic difference in gains to intermarriage of $EU10$ and $EU2$ countries with respect to all other countries, remains constant over time and not statistically significant. Hence, differences in gains to intermarriage with East European women are affected by legal status acquisition only after the EU enlargements, but not before. Secondly, the change in legal status induces a sharp negative effect on gains to intermarriage for $EU10$ after the announcement of the first EU enlargement, in line with anticipatory effects, and an even larger decrease afterwards, which remained persistent in the long run. Finally, estimates uncover a particular evolution of the β_t coefficients for $EU2$ countries. They show a substitution effect in the propensity to intermarry between $EU10$ and $EU2$ countries, i.e. between East European countries that enter the union during the first enlargement and those that were admitted with the second enlargement. For instance, after the announcement of the first EU enlargement, $EU2$ countries experience a positive change in the gains to intermarriage, followed by a sharp negative and persistent effect at the moment of the second enlargement.

3.6.3 Heterogeneous Effects across Marriage Markets

So far, I document that the acquisition of the legal status by foreign women of new EU member countries induced a decrease in their gains to intermarriage. The decrease is heterogeneous and especially strong for $EU2$ minorities that were less integrated from a socio-economic point of view. For those minorities, intermarrying might represent a valuable alternative to lacking or costly labor market opportunities.

In this section, I test for the potential heterogeneity in the effect of legal status across marriage markets. Across-markets heterogeneity is consistent with the hypothesis that the acquisition of the legal status might have changed the propensity to intermarry, especially in those markets where the access to labor opportunities for legalized migrants is easier or less costly. I proxy the heterogeneity in labor opportunities across markets, exploiting variability in migrants share by political-ethnic group. I argue that markets that are more

intensely populated by migrants at the same time are markets with higher labor market opportunities, in accordance with a large literature documenting that the residential segregation of migrants is driven by ethnic-labor network opportunities [f.e. Borjas (1995)]. As reported in Figure 3.5, the migrants' residential distribution across provinces displays a massive economic North-South divide, in line with Mastrobuoni and Pinotti (2015). To study the heterogeneity in the effect of legal status acquisition, I construct a binary variable, $Share\_Migrants_j$, equal to one for provinces with a migration rate larger than the Italian mean, by ethnic-group j, and zero otherwise [19].

To rule out potential endogenous increase in migration inflows by ethnic-group and province, in response to the EU enlargements, I exploit preexisting variation in migration intensity by ethnic-group across markets. For instance, I provide evidence of a stable spatial distribution of migrants residential selection over time. Figure 3.5 shows that provinces with higher migration intensity in 1995 also had higher migration intensity in 2012, pointing to a strong persistence in residential segregation, before and after the EU enlargements. I investigate the across-markets heterogeneity in legal status acquisition effects, comparing markets with low migration intensity as opposed to markets with high migration intensity, via a triple DID estimation. For instance, I augment the model in (3.8), interacting the variables with the dummy $Share\_Migrants_j$. If legal status reduces intermarriage gains by enhancing the economic-labor prospects for legalized migrants, I could expect the decrease to be larger in more migration-attractive markets. Estimates are reported in Table 3.3. Results are consistent with the hypothesis described. Indeed, the variation in the effect of the legal status acquisition on the gains to intermarriage for $EU10$ and $EU2$ is larger in markets with larger migration intensity, of 16.8% to 44.2%. Estimates are statistically significant, accounting for anticipatory affects.

3.7 Conclusions and Future Directions

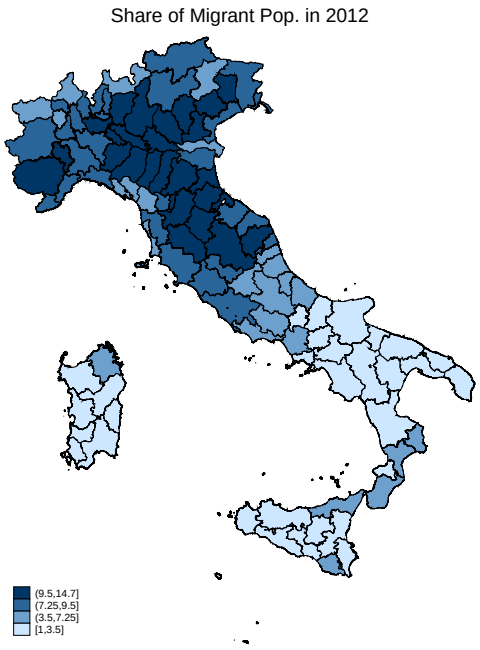
This research proposes and estimates a transferable utility marital matching model along the cultural-ethnic identity of spouses. I argue that gains to interethnic marriage depend on both cultural preferences and legal status or citizenship acquisition motives. The analysis focuses on interethnic unions, between one Italian spouse and one non-Italian spouse. Following Choo and Siow (2006b)'s contribution, I estimate the gains to intermarriage, thanks to administrative individual level data on marriages formed in Italy from 1997 to 2012, and individual Census data on population vectors. I estimate the price of citizenship, exploiting exogenous variation in the legal status acquisition driven by the EU enlargements to

[19] Data on the population distribution by ethnicity and by province come from municipality records on the movements of the foreign resident population. Population shares by ethnicity might be calculated thanks to municipality data on the total resident population for the same time interval.

East European countries in 2004 and 2007. Hence, I estimate the effect of legal status acquisition via a DID strategy. I compare the relative change in gains to intermarriage in the post-enlargement periods relative to the pre-enlargements one, between new EU member countries and countries for whom the legal status regulation did not change with the EU enlargements. I document that the legal status or citizenship acquisition affects marital matching choices, reducing intermarriage incentives. The analysis explicitly accounts for spillover effects generated in the market in terms of cross-ethnic marital substitutability, exploiting different counterfactual groups.

In the future, I would like to investigate deeply substitution patterns in response to exogenous variation in legal status acquisition. In particular, I plan to work in this direction proposing a multidimensional matching framework where I increase the number of spouses relevant attributes for marital decisions, in terms of age and education. Preliminary evidence on age and educational assortative is reported in Appendix C.1.1.

Figure 3.5: Distribution of Migrant Population by Province, 2012



Note: *Sources:* Movements of the foreign resident population (2012), Italy. The map reports the percentage share of migrant population over the total resident population by province in 2012. The color classification corresponds to the quartiles of the population distribution.

Chapter 4

Ethnic Judicial Bias: Discrimination or Integration? Evidence from Separation and Divorce Sentences

Abstract

This project aims to understand whether judicial decisions respond to the ethnic identity of spouses and what incentives those judgements are guided, by looking at family law proceedings. I analyse judicial decisions from the universe of separation and divorce sentences in Italy, from 2000 to 2012.Studying the legal custody assignment of children, I have documented that it is not the mother ethnic identity per se that affects her custody probability, but a significant differential is detected interacting mothers' ethnicities with the family type. For instance, compared to native mothers in homogeneous marriages, foreign mothers in mixed families are significantly less likely to obtain their children custody, while foreign mothers in homogeneous families are favored. Estimates are heterogeneous across ethnic groups. The project takes advantage of the random assignment of cases to judges to rule out potential unobserved case heterogeneity and it will provide between-judges evidence of ethnic judicial disparities. New original data from transcripts of judges decisions, from the Family section of the Court of Milan, allows to improve the research, both methodologically and conceptually, in the ability to discriminate across different potential motivations that drive judges' sentences.

Keywords: Ethnic Bias, Separation and Divorce Sentences, Custody Probability.
JEL Classification: J12, J15, J16, K36.

4.1 Introduction

The third article of the Italian Constitution establishes that all citizens are equal at law, without distinction of sex, race, language, religion, political opinions, personal and social conditions. Despite several countries explicitly in their Constitutions state that all citizens have equal judicial rights, huge disparities in the application of the law have been called into question. A special attention has been devoted to settings where cultural and ethnic barriers are particularly pronounced, as in Israel where religious differences between Arabs and Jews determine strong social identification and ingroup bias, or settings where members of minority groups are likely to be sentenced by juries not belonging to the same minority. The black-white divide in the US has drawn this kind of attention repeatedly, as blacks are largely segregated. Motivated by anecdotal evidence of cases of judicial discrimination together with statistical evidence of unequal judicial outcome, several economic and legal studies have investigated and documented the existence of ethnic and racial bias in different jurisdictions, e.g. civil litigations between private parties (Shayo and Zussman (2011)), capital sentences (Alesina and La Ferrara (2014)), and criminal trials (Anwar et al. (2012), Abrams et al. (2012)).

Using judiciary data about legal separations and divorces, in this project I aim to understand whether judicial decisions respond to the ethnic identity of spouses and what incentives those judgements are guided by. Specifically, I analyse judicial decisions from separation and divorce sentences in Italy, between 2000 and 2012.

This setting is particularly interesting for several reasons. First, by analysing the role of ethnic identity on family law dispositions, this project contributes to the literature on judicial bias in a perspective that has been greatly overlooked. As a matter of fact, a large body of evidence has shown the role of post-marital arrangements in general, and of child custody assignments, in particular, in shaping long-term educational and economic outcomes (e.g. Del Boca and Flinn (1995) and Dahl and Moretti (2008), among many others). Hence, in this project, I study the child custody assignment, as a crucial and relevant phase, especially in early childhood, for the child socio-economic development. Secondly, compared to previous studies concerned with racial or religious asymmetries, this project exploits ethnic variability across migrants within Italy, in accordance with the continuing increase of migratory phenomena. Indeed, the period under investigation is characterized by intense migratory inflows, making ethnic boundaries more and more salient, as compared, for example, to religious ones. Because migration rises different issues concerning the integration of migrants into new communities, a deep investigation of the mechanisms that promote or obstacle such integration is of primary necessity. Another peculiar characteristic of the Italian judicial system is that, within courts, proceedings are randomly assigned to

judges[1]. This way I can take advantage of this randomization process to rule out potential unobserved case heterogeneity in the estimation of the ethnic bias. Finally, I am currently collecting original data from transcripts of judges decisions, from the Family section of the Court of Milan, for the same period. Information embodied in these documents allows to improve the research, both methodologically and conceptually, in the ability to discriminate across different potential motivations driving judges' sentences.

The empirical analysis is based on unique and high-quality administrative data on separation and divorce records, collected from the registries of civil court chancelleries, and provided by ISTAT (National Statistical Institute). The datasets cover the universe of separations and divorces registered in Italy in the period from 2000 to 2012. They provide information on (i) the personal characteristics of the spouses at marriage, (ii) retrospective information on marriage and family characteristics, (iii) information on the separation and divorce proceedings, and (iv) post-dissolution arrangements, such as alimony obligations, recipient subject and annual amount of alimony, the assignment of child custody and the disposition of the family home. The analysis focuses on 525,964 separations, 10.38 percent of which involve a foreign parent and 222,680 divorces, with a 10 percent of foreign parents.

By looking at the probability of *legal* custody assignment to the mother, I provide evidence that mothers' ethnic identity has no effect per se, while a large significant differential is detected interacting mothers' ethnic identity with father ethnic identity, exploiting a difference-in-differences strategy. For instance, compared to native mothers in homogeneous marriages, foreign mothers in mixed families are significantly less likely to obtain their children custody by 1.4 to 3.5 percentage points, while foreign mothers in homogeneous families are favored, with a significant increase in custody probability of about 8.2 to 11.2 percentage points, in separation and divorces respectively. Results are heterogeneous across ethnic groups. Building on these results, I extend the empirical evidence to the probability of *effective* custody assignments to the mother, considering additional post-dissolution outcomes, such as alimony obligations and family home disposition, providing evidence of a similar differential pattern.

What is challenging in this setting is to identify ethnic judicial bias in presence of systematic ethnic differences in both observable and unobservable characteristics. I rule out potential confounding in unobserved case heterogeneity, including court fixed effects, relying on the fact that within each court judge assignment is orthogonal to parents' characteristics. On top of that, I account for heterogeneity in mothers and father characteristics, by including a large set of observable dimensions at the individual and household level, such as mother and father educational attainment, age and past marital history; occupation and labor market controls and province population size, as well as year fixed effects to con-

[1]Refer to Coviello et al. (2015) for a detailed discussion on the randomization process.

trol for systematic trends in maternal custody assignment over time. Despite the rich set of observable dimensions, I cannot exclude the possibility that foreign mothers differ with respect to native ones in some relevant unobserved dimensions.

I plan to overcome the endogeneity issue, enriching my analysis with new original data. I am currently collecting rich data on transcripts of judge decisions, from the Milan Family Court for the same period 2000-2012. In part, transcripts are electronically stored, while the most part is stored in the hardcopy archive. The data collection requires a long process of search, selection and scan of transcripts. By using modern text analysis techniques on judgement transcripts, I can extract information (among others) on the judge responsible for each proceeding, the overall history of the proceeding, the intervention and report of third parties whether involved, and the motivations that drive the final judges' decisions, especially in judicial proceedings.

Importantly, new information will allow me to expand the scope of the research along several directions, estimating the extent of ethnic judicial bias across judges, but also exploring its sources. As a first step, I plan to link judges data to corresponding proceedings. By augmenting my baseline specification with judge fixed effects, I can identify the presence of ethnic discrepancies between-judges in family law sentencing. Borrowing from Abrams et al. (2012), I rely on the random assignment of cases to judges to control for potential omitted variable bias. Because of the randomization process, process which ensures that systematic differences in observable and unobservable characteristics of cases by spouses ethnic group are orthogonal across judges, any heterogeneity across judges in the child custody assignment to the mother is driven by judges' differential responses to the mother ethnic identity [2] As a second step, I would like to dig more into judicial incentives. In particular, I consider trying disentangling two main channels: child integration incentives as opposed to foreign mothers bias, driven by potential out-migration. On the one hand, judges might have an incentive to maximize the likelihood of integration of minors from heterogeneous families to the host society, granting the custody to the native spouse. Indeed, the cultural transmission channel might be especially binding for mothers belonging to minorities, which are culturally far from the host country, or with lower integration probability in terms of employment opportunities. On the other hand, judges might have the incentive to minimize the out-migration probability of children, granting the larger contact possible with both parents, especially with the non-custodial one. Thus, I expect the role of incentives to vary across mothers' ethnicities.

In the following section, I provide an account of the existing literature on racial and

[2] The random assignment of cases to judges has been exploited in several settings to identify, for example, the black-white differential in incarceration rates (Abrams et al. (2012) and Manudeep et al. (2016)), the differential in disability insurance (Autor et al. (2015)) and the impact of the intergenerational transmission of welfare attitudes (Dahl et al. (2014).

ethnic judicial bias. In Section 4.2 the data is introduced. Section 4.3 outlines the empirical strategy to investigate the judicial response to spouses' ethnic identity, while Section 4.3.3 discusses potential incentives driving judgements. Finally, Section 3.7 concludes.

4.2 The Data

The empirical analysis is based on unique and high-quality administrative data on separation and divorce records, collected from the registries of civil court chancelleries, and provided by ISTAT (National Statistical Institute). The datasets cover the universe of separations and divorces registered in Italy in the period between 2000 to 2012[3]. They provide information on (i) the personal characteristics of the spouses at marriage, such as age, educational attainment, professional status and occupation, citizenship status and country of origin for foreign-born individuals, (ii) retrospective information on marriage and family characteristics such as the date of the marriage, the province of residence, the number of children born in the marriage and the demographic characteristics of minor children present in the household, (iii) information on the separation and divorce proceedings in terms of filing and disposition date, type of procedure (consensual or judicial), type of conclusion of the proceeding (conciliation, approval, cancellation, court acceptance, judge remittance or change of proceeding), and (iv) post-dissolution arrangements, such as alimony obligations, recipient subject and annual amount of alimony, the assignment of child custody and the disposition of the family home. Proceedings classified to end with conciliation, cancellation, or change of rite are registered, but no information is available for them. They have been dropped because not representative of effective marital dissolution choices. The final sample counts 525,964 separations, 10.38 percent of which involve a foreign parent and 222,680 divorces, with a 10 percent of foreign parents.

The study of the Italian context is particularly interesting, for two main reasons. First, Italy legalized divorce only quite recently compared with other Western countries. Thus the evaluation of the marriage dissolution pattern over the last two decades is particularly relevant. Divorce was legalized in Italy by Law 898/1970. After 1970, only two major reforms intervene to modify it. The first reform (Law 436/1978) improves the economic protection of the weaker partner guaranteeing heath insurance assistance. The second reform (Law 74/1987) lowers the separation period requested before a divorce application

[3] The microdata about separations and divorces comes from questionnaires completed directly by the courts' chancelleries. The fact that information comes from objective sources (i.e. information does not come from individual self-reporting or retrospective acknowledge) increases the reliability of the analysis. For the period under investigation, registries of civil court chancelleries constitute a unique source for separations and divorces data, while starting from December 2014 (in application of Law 162/2014) consensual separation and divorce proceedings can be submitted to the civic registrar. This excludes potential issues in the selection of available data.

(from 5 to 3 years). Changes in the legal structure of separation and divorce have been shown to increase the marital dissolution incidence [Friedberg (1998)]. During the period under investigation, only one relevant divorce law took place in 2006 (Law 54/2006), which changed the standard arrangement from sole custody to joint custody of minor children. A detailed description of the law provisions is presented by De Blasio and Vuri (2013).

I am currently collecting original data on transcripts of judge decisions, from the Milan Family Court for the same period 2000-2012. Part of the transcripts are electronically stored, while the most part is stored in the hardcopy archive. By using modern text analysis techniques on judgement transcripts, I can extract information (among others) on the judge responsible for each proceeding, the overall history of the proceeding, the intervention and report of third parties whether involved, and the motivations that drive the final judges' sentences, especially in judicial proceedings. Importantly, I can perfectly match information from transcripts with ISTAT data, exploiting the filing and disposition date of the proceedings.

4.3 Judicial Response to Spouses' Ethnic Identity

4.3.1 Estimation Strategy and Empirical Evidence

The empirical analysis exploits three different measures of custody outcome. The main outcome is a binary variable that takes the value of 1 in case of *sole legal* mother custody of the first child. Moreover, in consideration of the introduction of the reform (Law 54/2006) that changed the default arrangement from sole custody to joint custody of minor children, I also consider the *effective* probability of mother custody of the first child. I recover the *effective* probability in two different ways. First, I consider as outcome the probability of *effective* mother custody of the first child where in case of legal joint custody the criterion of the home destination is exploited. Secondly, I consider as outcome the probability of *effective* mother custody of the first child, where the criterion of the indirect child support measure is exploited (i.e. the indirect child support measure the probability that the father is obliged to provide monetary resources for child maintenance to the mother, De Blasio and Vuri (2013)).

Consider the descriptive statistics for the outcome variables of interest, conditioning on mother and father ethnic identity. By looking at the sole mother legal custody probability, native mothers are more likely to obtain their children custody compared to foreign mothers. The pattern is similar for separation (Panel A) and divorce (Panel B) proceedings. This in itself does not necessary provide evidence in favour of ethnic judicial bias, because native mothers might be systematically different from foreign mothers in some relevant observable and unobservable dimension. However, if this was the only reason underlying

the difference, I would expect to observe a similar outcome probability regardless of father ethnicity. In fact, conditioning on the mother ethnicity, a foreign mother is 2.3 percentage points more likely to obtain the custody of children upon separation if married with a foreign spouse compared to a marriage with a native spouse. The same difference increases to 7.2 percentage points in case of divorce. This suggest that judges choices are heterogeneous interacting mothers' ethnic identity with father ethnic identity.

Figure 4.1 displays this pattern graphically, by reporting the probability of sole mother legal custody by mother and father ethnicity. What is particularly interesting is the differential judges' response to the mother ethnic identity, as a function of the ethnic identity of the father, i.e. the differential change in the sole mother legal custody probability between native and foreign mothers, married with a native compared to a foreign spouse. Unconditional estimates reported in the figure uncover a mild positive differential change in separation data of about 0.30 percentage points and a larger differential change of 2.3 percentage points in divorce data.

I further investigate this differential pattern in judicial decisions, estimating the following econometric model.

$$\begin{aligned} y_{ij,ct} =& \alpha_0 + \alpha_1 MotherForeign_i + \alpha_2 FatherForeign_j \\ &+ \alpha_3 MotherForeign_i \times FatherForeign_j + \gamma_c + \theta_t + \varepsilon_{ij,ct} \end{aligned} \tag{4.1}$$

where y_{ijct} is the outcome variable, i.e. the probability of sole legal custody assignment to the mother i, married with father j, living in a province under court jurisdiction c, at time t. $MotherForeign_i$, $FatherForeign_j$, and the interaction term $MotherForeign_i \times FatherForeign_j$ are indicator variables, denoting whether the mother, the father or both are foreign-born.

In addition, I include both year fixed effects, θ_t, and court fixed effects, γ_c. Year fixed effects allow to control for systematic trends in maternal custody assignment over time. By including court fixed effects, I control for systematic differences across courts in maternal custody assignment, potentially driven by systematic selection in residential location. On top of that, courts fixed effects allow to rule out potential confounding in unobserved case heterogeneity, relying on the fact that within each court judge assignment is orthogonal to parents' characteristics. For instance, as described by Coviello et al. (2015), the cases filed during the day are assigned to judges following their alphabetic name ordering, starting from a randomly extracted letter. Finally, $\varepsilon_{ij,ct}$ is an error term clustered at the court level. There are 165 courts in the sample. Results from separation proceedings are reported in Table 4.2 in Columns (1-3), where I progressively add the indicator variables, $MotherForeign_i$, $FatherForeign_j$, and $MotherForeign_i \times FatherForeign_j$. Results from divorce proceedings are

reported in Table 4.3, in Columns (1-3).

By looking at the probability of *legal* custody assignments to the mother, I provide evidence that mothers' ethnic identity has no effect per se as reported in Columns 1 and 2, while a large significant differential is detected interacting mothers' ethnic identity with father ethnic identity, exploiting a difference-in-differences strategy. For instance, compared to native mothers in homogeneous marriages, foreign mothers in mixed families are significantly less likely to obtain their children custody by 1.3 (3.6) percentage points upon separation (divorce), while foreign mothers in homogeneous families are favored, with a significant increase in custody probability of about 8.8 (11.4) percentage points.

What is challenging in this setting is to identify ethnic judicial bias in presence of systematic ethnic differences in both observable and unobservable characteristics, relevant for custody assignment. I augment the model in (4.1), including a set of individual and household observable characteristics potentially affecting the outcome, as follows:

$$\begin{aligned} y_{ij,ct} =& \alpha_0 + \alpha_1 MotherForeign_i + \alpha_2 FatherForeign_j \\ &+ \alpha_3 MotherForeign_i \times FatherForeign_j + \mathbf{X}'_{ij}\beta + \gamma_c + \theta_t + \varepsilon_{ij,ct} \end{aligned} \tag{4.2}$$

where $\mathbf{X}'_{ij}$ accounts for heterogeneity in mothers and father characteristics at the individual and household level, such as mother and father educational attainment, mother and father age and mother and marital history, mother and father occupation and labor market controls and province population size. In Columns (4) and (6) of Table 4.2 and Table 4.3, I progressively add these sets of controls. Estimates of the main explanatory variables are robust to the inclusion of the additional controls. In a marriage with a native husband, foreign mothers are significantly less likely to obtain their children custody by 1.3 (3.7) percentage points compared to native mothers, upon separation (divorce). In addition, in a marriage with a foreign spouse, native mothers are significantly more likely to obtain their children custody by 4.9 (6.3) percentage points compared to marriage with a native spouse. Finally, the differential change in the sole mother legal custody probability between foreign and native mothers, married with a foreign compared to a native spouse is positive and statistically significant.

I further investigate the differential pattern in judicial decisions across the ethnic identity of spouses, examining two subsamples. I distinguish between consensual and judicial proceedings. Indeed, in judicial proceedings, judges have an active role in deciding upon post-marital dissolution agreements and children custody assignment. On the contrary, in consensual proceedings, judges approve spouses' arrangements, ensuring their compliance with the law. This contrast might potentially generate differences in the extent of ethnic judicial bias. Estimates remain consistent with the above discussion. To notice that estimates

are stronger in magnitude for the subsample of judicial proceedings, for both separation and divorce cases.

Building on these results, I extend the empirical evidence to the probability of *effective* custody assignments to the mother, considering additional post-dissolution outcomes, such as family home dispositions and alimony obligations, providing evidence of a similar differential pattern in judicial response to spouses ethnic identity.

4.3.2 *Heterogeneity in Judges Responses to Spouses' Ethnic Identity*

Despite the rich set of observable dimensions, I cannot exclude the possibility that foreign mothers and fathers differ with respect to native ones in some relevant unobserved dimensions.

I plan to overcome the endogeneity issue, thanks to the new data collection, described in Section 4.2. For instance, I will link judges data to corresponding proceedings. By augmenting my model specification in (4.2) with judge fixed effects, I can identify the presence of ethnic discrepancies between-judges in family law sentencing. Borrowing from Abrams et al. (2012), I rely on the random assignment of cases to judges to control for potential omitted variable bias. Because of the random assignment of cases, i.e. systematic differences in observable and unobservable characteristics of cases by spouses ethnic group are orthogonal across judges, any heterogeneity across judges in the child custody assignment is driven by a differential response to mothers' ethnic identities [4].

4.3.3 *Understanding Judicial Incentives*

Moreover, I would like to dig more into judicial incentives. In particular, I consider to disentangle two main channels: child integration incentives from foreign mothers bias (driven by a potential increase in the likelihood of out-migration), using modern text analysis techniques on judgements transcripts. On the one hand, judges might have an incentive to maximize the likelihood of integration of minors from heterogeneous families within the host society, granting the custody assignment to the native spouse. The cultural transmission channel is relevant, especially if mothers belongs to minorities far from the host country, or with lower integration probability in terms of employment opportunities. On the other hand, judges might have the incentive to minimize the out-migration probability of children, granting the larger contact possible with both parents, especially with the one not assigned to custody. I expect the role of incentives to vary across mothers' ethnicities.

[4] The random assignment of cases to judges has been exploited in several settings to identify, for example, the black-white differential in incarceration rates (Abrams et al. (2012) and Manudeep et al. (2016)), the differential in disability insurance (Autor et al. (2015)) and the impact of the intergenerational transmission of welfare attitudes (Dahl et al. (2014).

4.4 Conclusions and Future Directions

The present research aims to understand whether judicial decisions respond to the ethnic identity of spouses and what incentives those judgements are guided, by looking at family law proceedings. I analyse judicial decisions from the universe of separation and divorce sentences in Italy, between 2000 and 2012. Studying the legal custody assignment of children, I have documented that it is not the mother ethnic identity per se that affects her custody probability, but a significant differential is detected interacting mothers' ethnicities with the family type. For instance, compared to native mothers in homogeneous marriages, foreign mothers in mixed families are significantly less likely to obtain their children custody, while foreign mothers in homogeneous families are favored. Estimates are heterogeneous across ethnic groups. The project takes advantage of the random assignment of cases to judges to rule out potential unobserved case heterogeneity and it will provide between-judges evidence of ethnic judicial disparities.

New original data from transcripts of judges decisions, which I am currently collecting from the Family section of the Court of Milan, allows to improve the research, both methodologically and conceptually, in the ability to discriminate across different potential motivations that drive judges' sentences. In particular. I would like to disentangle child integration incentives from foreign mothers bias (driven by a potential increase in the likelihood of out-migration), using modern text analysis techniques on judgements records. I expect the role of incentives to vary across mothers' ethnicities.

Bibliography

Abrams, D. S., M. Bertrand, and S. Mullainathan (2012). Do Judges Vary in Their Treatment of Race? *The Journal of Legal Studies 41*(2), 347–383.

Alesina, A. and E. La Ferrara (2014). A Test of Racial Bias in Capital Sentencing. *American Economic Review 104*(11), 3397–3433.

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

Amato, P. R. (1993). Children's Adjustment to Divorce: Theories, Hypotheses, and Empirical Support. *Journal of Marriage and the Family 55*(1), 23–38.

Amato, P. R. (1996). Explaining the Intergenerational Transmission of Divorce. *Journal of Marriage and the Family 58*(3), 628–640.

Amato, P. R. (2000). The Consequences of Divorce for Adults and Children. *Journal of Marriage and Family 62*(4), 1269–1287.

Amuedo-Dorantes, C., C. Bansak, and S. Raphael (2007). Gender Differences in the Labor Market: Impact of IRCA's Amnesty Provisions. *The American Economic Review 97*(2), 412–416.

Anwar, S., P. Bayer, and R. Hjalmarsson (2012). The Impart of Jury Race in Criminal Trials. *Quarterly Journal of Economics 127*, 1017–1055.

Autor, D., A. R. Kostol, and M. Mogstad (2015). Disability Benefits, Consumption Insurance, and Household Labor Supply. Technical report.

Azzolini, D., R. Guetto, et al. (2015). The Impact of Citizenship on Intermarriages. Quasi-Experimental Evidence From Two European Union Eastern Enlargements. Technical report, Research Institute for the Evaluation of Public Policies (IRVAPP), Bruno Kessler Foundation.

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

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

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

Becker, G. S. (1991). *A Treatise on the Family*. Cambridge, Massachusetts: Harvard University Press.

Becker, G. S., E. M. Landes, and R. T. Michael (1977). An Economic Analysis of Marital Instability. *The Journal of Political Economy 85*(6), 1141–1187.

Becker, G. S. and C. B. Mulligan (1997). The Endogenous Determination of Time Preference. *Quarterly Journal of Economics*, 729–758.

Bisin, A., P. Eleonora, V. Thierry, and Z. Yves (2016). Bend it like Beckham: Ethnic Identity and Integration. *European Economic Review 90*, 146–164.

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.

Blundell, R., P.-A. Chiappori, T. Magnac, and C. Meghir (2007). Collective Labour Supply: Heterogeneity and Non-Participation. *The Review of Economic Studies 74*(2), 417–445.

Borjas, G. J. (1995). Ethnicity, Neighborhoods, and Human-Capital Externalities. *American Economic Review 85*(3), 365–90.

Boyd, R. (1988). *Culture and the Evolutionary Process*. University of Chicago Press.

Bratter, J. L. and R. B. King (2008). But Will It Last?: Marital Instability Among Interracial and Same-Race Couples. *Family Relations 57*(2), 160–171.

Breslow, N. E. (1974). Covariance Analysis of Censored Survival Data. *Biometrics 30*(1), 89–99.

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

Bruze, G., M. Svarer, and Y. Weiss (2015). The Dynamics of Marriage and Divorce. *Journal of Labor Economics 33*(1), 123–170.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Cavalli-Sforza, L. L. (1981). *Cultural Transmission and Evolution: a Quantitative Approach*. Number 16. Princeton University Press.

Cavalli-Sforza, Luigi, P. M. and A. Piazza (1994). *The History and Geography of Human Genes*. Princeton University Press.

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 (2008). An assignment model with divorce and remarriage. Technical report.

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., S. Oreffice, and C. Quintana-Domeque. Bidimensional Matching with Heterogeneous Preferences: Smoking in the Marriage Market. *Journal of the European Economic Association. Forthcoming*.

Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque (2012). Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market. *Journal of Political Economy 120*(4), 659–695.

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

Chiappori, P.-A., B. Salanié, and Y. Weiss. Partner Choice and the Marital College Premium: Analyzing Marital Patterns Over Several Decades. *American Economic Review. Forthcoming*.

Chiswick, B. R. and P. W. Miller (2014). International migration and the economics of language. *Handbook of the Economics of Immigration*, 211–269.

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

Choo, E. and A. Siow (2006a). Estimating a marriage matching model with spillover effects. *Demography 43*(3), 463–490.

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

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

Constant, A. F., L. Gataullina, and K. F. Zimmermann (2009). Ethnosizing Immigrants. *Journal of Economic Behavior & Organization 69*(3), 274–287.

Constant, A. F. and K. F. Zimmermann (2008). Measuring Ethnic Identity and Its Impact on Economic Behavior. *Journal of the European Economic Association 6*(2-3), 424–433.

Coviello, D., A. Ichino, and N. Persico (2015). The Inefficiency of Worker Time Use. *Journal of the European Economic Association 13*(5), 904–947.

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 187–220.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014). Family welfare cultures. *Quarterly Journal of Economics 129*(4), 1711–1752.

Dahl, G. B. and E. Moretti (2008). The Demand for Sons. *The Review of Economic Studies 75*, 1085–1120.

Dalla Zuanna, G. (2008). A Marriage-Cohort Analysis of Legal Separations in Italy. *Population 63*(1), 173–193.

De Blasio, G. and D. Vuri (2013). Joint Custody in the Italian Courts.

De Rose, A. (1992). Socio-Economic Factors and Family Size as Determinants of Marital Dissolution in Italy. *European Sociological Review 8*(1), 71–91.

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

Del Boca, D. and C. J. Flinn (1995). Rationalizing Child-Support Decisions. *American Economic Review 85*(5), 1241–1262.

Doepke, M. (2015). Gary Becker on the Quantity and Quality of Children. *Journal of Demographic Economics 81*(01), 59–66.

Doepke, M. and F. Zilibotti (2008). Occupational Choice and the Spirit of Capitalism. *Quarterly Journal of Economics 123*(2), 747–793.

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

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

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

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.

Efron, B. (1977). The Efficiency of Cox's Likelihood Function for Censored Data. *Journal of the American Statistical Association 72*(359), 557–565.

Egger, P. H. and F. Toubal (2016). Common Spoken Languages and International Trade. In *The Palgrave Handbook of Economics and Language*, pp. 263–289. Springer.

Fernández, R. (2013). Cultural Change as Learning: The Evolution of Female Labor Force Participation Over a Century. *American Economic Review 103*(1), 472–500.

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.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *Quarterly Journal of Economics 119*(4), 1249–1299.

Fernández, R., N. Guner, and J. Knowles (2005). Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality. *Quarterly Journal of Economics 120*(1), 273–344.

Fernández, R. and J. Wong (2014). Unilateral Divorce, the Decreasing Gender Gap, and Married Women's Labor Force Participation. *American Economic Review 104*(5), 342–347.

Ferro, I. and D. Vignoli (2009). Rising Marital Disruption in Italy and Its Correlates. *Demographic Research, Max Planck Institute for Demographic Research, Rostock, Germany 20*(4), 11–36.

Figlio, D., P. Giuliano, U. Özek, and P. Sapienza (2016). Long-Term Orientation and Educational Performance. Technical report, National Bureau of Economic Research N. 22541.

Fouka, V. (2016). Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I. Technical report, Working Paper.

Friedberg, L. (1998). Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data. *American Economic Review 88*(3), 608–627.

Frimmel, W., M. Halla, and R. Winter-Ebmer (2013). Assortative Mating and Divorce: Evidence From Austrian Register Data. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 176*(4), 907–929.

Fryer, R. G. (2007). Guess Who's Been Coming to Dinner? Trends in Interracial Marriage Over The 20th Century. *The Journal of Economic Perspectives 21*(2), 71–90.

Fu, X. and T. B. Heaton (2008). Racial and educational homogamy: 1980 to 2000. *Sociological Perspectives 51*(4), 735–758.

Galichon, A. and B. Salanié (2015). Cupid's Invisible Hand: Social Surplus and Identification in Matching Models.

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

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014). Marry Your Like: Assortative Mating and Income Inequality. *American Economic Review 104*(5), 348–353.

Guiso, L., P. Sapienza, and L. Zingales (2009). Cultural Biases in Economic Exchange? *Quarterly Journal of Economics 124*(3), 1095–1131.

Hauk, E. and M. Saez-Marti (2002). On the Cultural Transmission of Corruption. *Journal of Economic Theory 107*(2), 311–335.

Hitsch, G. J., A. Hortaçsu, and D. Ariely (2010). Matching and Sorting in Online Dating. *American Economic Review 100*(1), 130–163.

Huang, F., G. Z. Jin, and L. C. Xu (2012). Love and money by parental matchmaking: Evidence from urban couples in china. *American Economic Review 102*(3), 555–560.

Jones, F. L. (1996). Convergence and Divergence in Ethnic Divorce Patterns: A Research Note. *Journal of Marriage and the Family* (1), 213–218.

Kalmijn, M., P. M. de Graaf, and J. P. Janssen (2005). Intermarriage and The Risk of Divorce in The Netherlands: The Effects of Differences in Religion and in Nationality, 1974–94. *Population Studies 59*(1), 71–85.

Kaplan, E. L. and P. Meier (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association 53*(282), 457–481.

Klasing, M. J. (2012). Cultural Change, Risk-Taking Behavior and the Course of Economic Development. *Risk-Taking Behavior and the Course of Economic Development (July 7, 2012)*.

Lehrer, E. L. (2008). Age at Marriage and Marital Instability: Revisiting the Becker–Landes–Michael Hypothesis. *Journal of Population Economics 21*(2), 463–484.

Lewis, M. P., G. F. Simons, and C. D. Fennig (2009). *Ethnologue: Languages of the World*, Volume 9. SIL International Dallas, TX.

Lozano, F. A. and T. Sorensen (2011). The Labor Market Value to Legal Status.

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

Manudeep, B., G. B. Dahl, K. V. Løken, M. Mogstad, et al. (2016). Incarceration, Recidivism and Employment. Technical report.

Mastrobuoni, G. and P. Pinotti (2015). Legal Status and the Criminal Activity of Immigrants. *American Economic Journal: Applied Economics 7*(2), 175–206.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105–142.

Melitz, J. and F. Toubal (2014a). Native Language, Spoken Language, Translation and Trade. *Journal of International Economics 93*(2), 351–363.

Melitz, J. and F. Toubal (2014b). Native Language, Spoken Language, Translation and Trade. *Journal of International Economics 93*(2), 351–363.

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

Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. *American Economic Review 102*(4), 1508–39.

Monahan, T. P. (1970). Are Interracial Marriages Really Less Stable? *Social Forces 48*(4), 461–473.

Peto, R. (1972). Contribution to the Discussion of the Paper by DR Cox. *Journal of the Royal Statistical Society, Series B 34*, 205–207.

Phillips, J. A. and M. M. Sweeney (2006). Can Differential Exposure to Risk Factors Explain Recent Racial and Ethnic Variation in Marital Disruption? *Social Science Research 35*(2), 409–434.

Pinotti, P. (2017). Clicking on Heaven's Door: The Effect of Immigrant Legalization on Crime. *The American Economic Review 107*(1), 138–168.

Porterfield, E. (1982). Black-American Intermarriage in the United States. *Marriage and Family Review 5*(1), 17–34.

Rosenfeld, M. J. (2008). Racial, Educational and Religious Endogamy in the United States: A Comparative Historical Perspective. *Social Forces 87*(1), 1–31.

Sáez-Martí, M. and A. Sjögren (2008). Peers and Culture. *The Scandinavian Journal of Economics 110*(1), 73–92.

Salvini, S. and D. Vignoli (2011). Things Change: Women's and Men's Marital Disruption Dynamics in Italy During a Time of Social Transformations, 1970–2003. *Demographic Research 24*(5), 145–174.

Schwartz, C. R. (2013). Trends and variation in assortative mating: Causes and consequences. *Annual Review of Sociology 39*, 451–470.

Shapley, L. S. and M. Shubik (1971). The assignment game i: The core. *International Journal of game theory 1*(1), 111–130.

Shayo, M. and A. Zussman (2011). Judicial Ingroup Bias in The Shadow of Terrorism. *The Quarterly Journal of Economics 126*(3), 1447–1484.

Shimer, R. and L. Smith (2000). Assortative matching and search. *Econometrica 68*(2), 343–369.

Siow, A. (2015). Testing Becker's Theory of Positive Assortative Matching. *Journal of Labor Economics 33*(2), 409–441.

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

Spolaore, E. and R. Wacziarg (2014). Fertility and Modernity. *Discussion Papers Series 779*.

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

Stevens, G. and R. Schoen (1988). Linguistic intermarriage in the united states. *Journal of Marriage and the Family 50*(1), 267–279.

Stevenson, B. and J. Wolfers (2006). Bargaining in the Shadow of the Law: Divorce Laws and Family Distress. *Quarterly Journal of Economics 121*(1), 267–288.

Stevenson, B. and J. Wolfers (2007). Marriage and Divorce: Changes and Their Driving Forces. *The Journal of Economic Perspectives 21*(2), 27–52.

Swadesh, M. (1952). Lexico-Statistic Dating of Prehistoric Ethnic Contacts: With Special Reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society 96*, 452–463.

Tabellini, G. (2008). The Scope of Cooperation: Values and Incentives. *Quarterly Journal of Economics 123*(3), 905–950.

Tabellini, G. (2010). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association 8*(4), 677–716.

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

Weiss, Y. and R. J. Willis (1985). Children as Collective Goods and Divorce Settlements. *Journal of Labor Economics 3*(3), 268–292.

Weiss, Y. and R. J. Willis (1997). Match Quality, New Information, and Marital Dissolution. *Journal of Labor Economics 15*(1), S293–S329.

Wong, L. Y. (2003). Why so only 5.5% of black men marry white women? *International Economic Review 44*(3), 803–826.

Yancey, G. and G. Yancy (2007). Experiencing Racism: Differences in the Experiences of Whites Married to Blacks and non-Black Racial Minorities. *Journal of Comparative Family Studies 38*(2), 197–213.

Zhang, Y. and J. Van Hook (2009). Marital Dissolution Among Interracial Couples. *Journal of Marriage and Family 71*(1), 95–107.

Appendices

Appendix A

Marital Formation and Dissolution of Interethnic Marriages. Evidence from Italian Data

A.1 Evolution of Marital Formation and Dissolution over Time

Recent contributions document a general and profound transformation in families over the last decades. Beyond the reduction in the marriage rate, the increase in cohabitation and the parallel increase in the number of children born out of wedlock; the increasing probability of family dissolution, as observed in Western countries, is perceived as the most striking and relevant modification [Stevenson and Wolfers (2007), Frimmel et al. (2013) and Browning et al. (2014)]. Figure A.1 shows the evolution of the crude marriage rate (per thousand inhabitants) in the major OECD countries in 1970, 1995 and 2010. In general, the Figure reveals a decreasing trend in marital formation. The average marriage rate in 1970 was 8.29% and decreased to 5.69% in 25 years. Despite most of the country's sharing a common pattern, there is considerable heterogeneity in terms of levels and growth rates. Similarly Figure A.2 reports the evolution of the crude divorce rate in OECD countries for the same years. The average divorce rate approximately duplicates in 25 years, passing from 1.22% in 1970 to 1.99% in 1995. As noted earlier, the graph highlights a significant heterogeneity among countries. For instance, in 2005 the divorce rate in Italy proxies at 1%, while United States register a divorce rate of 3.6% per thousand inhabitants, which is the lowest rate registered since 1970 [Stevenson and Wolfers (2007)].

Focusing on the Italian marriage market, Figure A.3 presents the series of the marriage, separation and divorce rate starting from 1970 when divorce was legalized by Law

898/1970 [1]. The evolution of the marriage rate follows a monotonically decreasing pattern, almost halving in 40 years from 7.5% at the begging of the series to 3.5% in 2012. On the other hand, Italian marital instability seems to be an exception, remaining particularly low in magnitude in respect of other Western countries. Despite this comparison, the increase in marital dissolution is the highest registered together with Spain among comparable countries [2]. As an example, in 1970 after the legalization of divorce, the divorce rate is equal to 0.3%, but then it doubles in one year to 0.6%, , and thereafter registers a growth rate of 2 points in the entire period under investigation. The separation rate highlights a similar increasing pattern over time. The main difference among the series relates to their level, i.e. the separation rate is systematically higher with respect to the divorce rate confirming that they are two different legal processes within the Italian institutional framework. In particular, in Italy the process leading to divorce requires a minimum period of separation of 3 years (it was reduced from 5 to 3 years in 1987). For instance, of the total number of legal separations registered in 1995, only 60% ended in a divorce in the following decade. The rest of the couples simply remain separated not to afford the cost a new legal formality [Salvini and Vignoli (2011)]. Evidence implies that the Italian marital dissolution probability is underestimated using the divorce rate and is better proxied by legal separation rate [Dalla Zuanna (2008)]. Following the literature, the evidence presented in the paper measures marital dissolution thorough separations and not divorces [De Rose (1992); Ferro and Vignoli (2009)]. The distance with respect to European countries' average is partially reduced taken into consideration regional variability, where social and cultural marked differences in family perception persist over time. On average, separations are more common in the North and in some Central regions, such as Tuscany and Lazio, while Southern regions display the lowest separation rates [3]. A moderate heterogeneity is evidenced for divorces, where all areas show a monotonically increasing pattern with different rates of increase. To notice that the crude rates of divorce in Southern regions nearly triplicate over the analysis time.

Different dimensions might explain the increasing trend in marital dissolution over time [Stevenson and Wolfers (2007)]. In general, the legal access to divorce becomes easier over time. The metamorphosis of social and legal norms increased the market value of women's time and strengthened the role of the state in regulating and enforcing monetary compensation to the disadvantaged spouse following divorce. These factors contributed to reduce

[1] The crude marriage, separation and divorce rate are defined as the absolute number of marriages, separations and divorces registered in a given year per thousand inhabitants. Statistics on marriages, separations and divorces come from ISTAT Data, Demographic Statistics (1970-2012). Annual data on the resident population in Italy (1995-1970) come from ISTAT elaborations on 1991, 2001 and 2011 Census data.

[2] Refer to Frimmel et al. (2013) and Browning et al. (2014) for a deep cross country comparison.

[3] Evidence of regional heterogeneity in separation and divorce decisions is reported in Figures A.4 by geographical macroarea.

the monetary costs of divorce, especially for women. On the other hand, the secularization process and the general increase in separations and divorces lowered the stigma of marital dissolution, thus reducing its social cost.

Figure A.1: Evolution of Crude Marriage Rates in OECD Countries

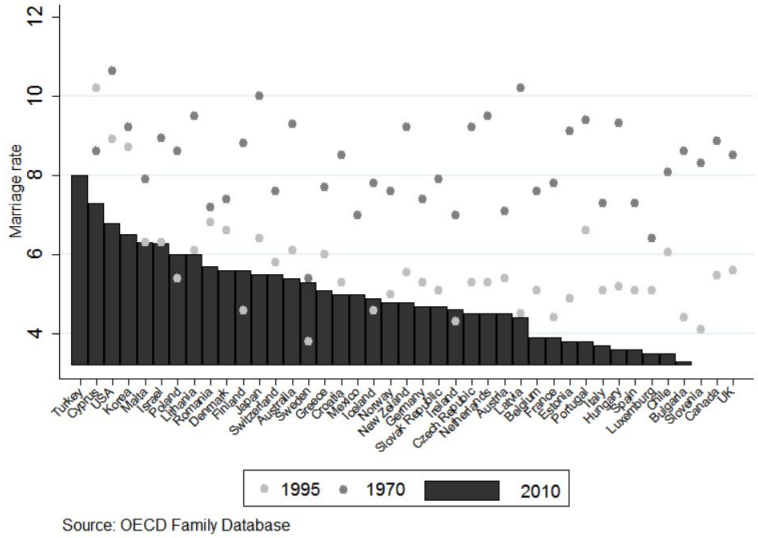
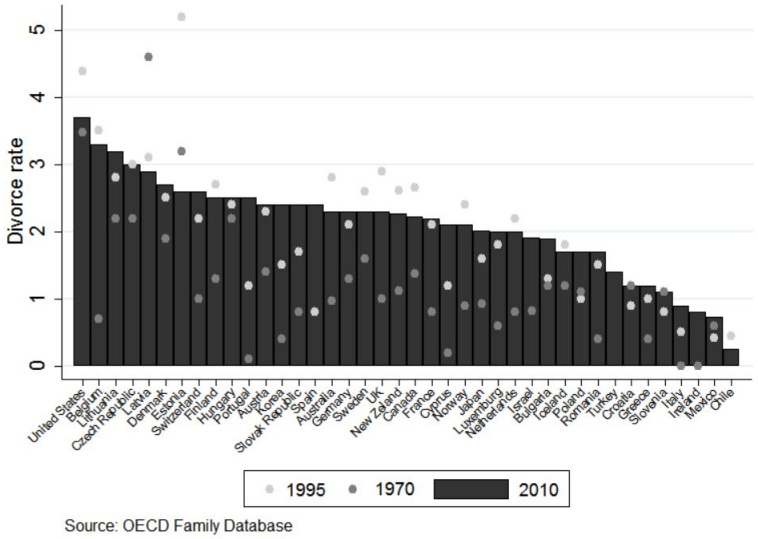
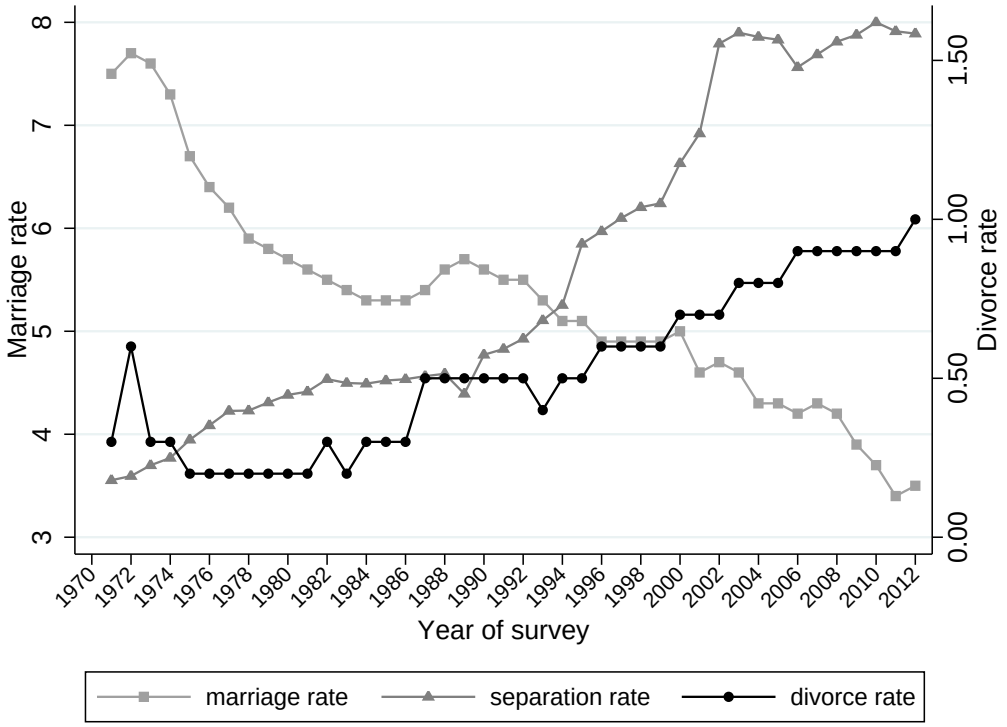


Figure A.2: Evolution of Crude Divorce Rates in OECD Countries



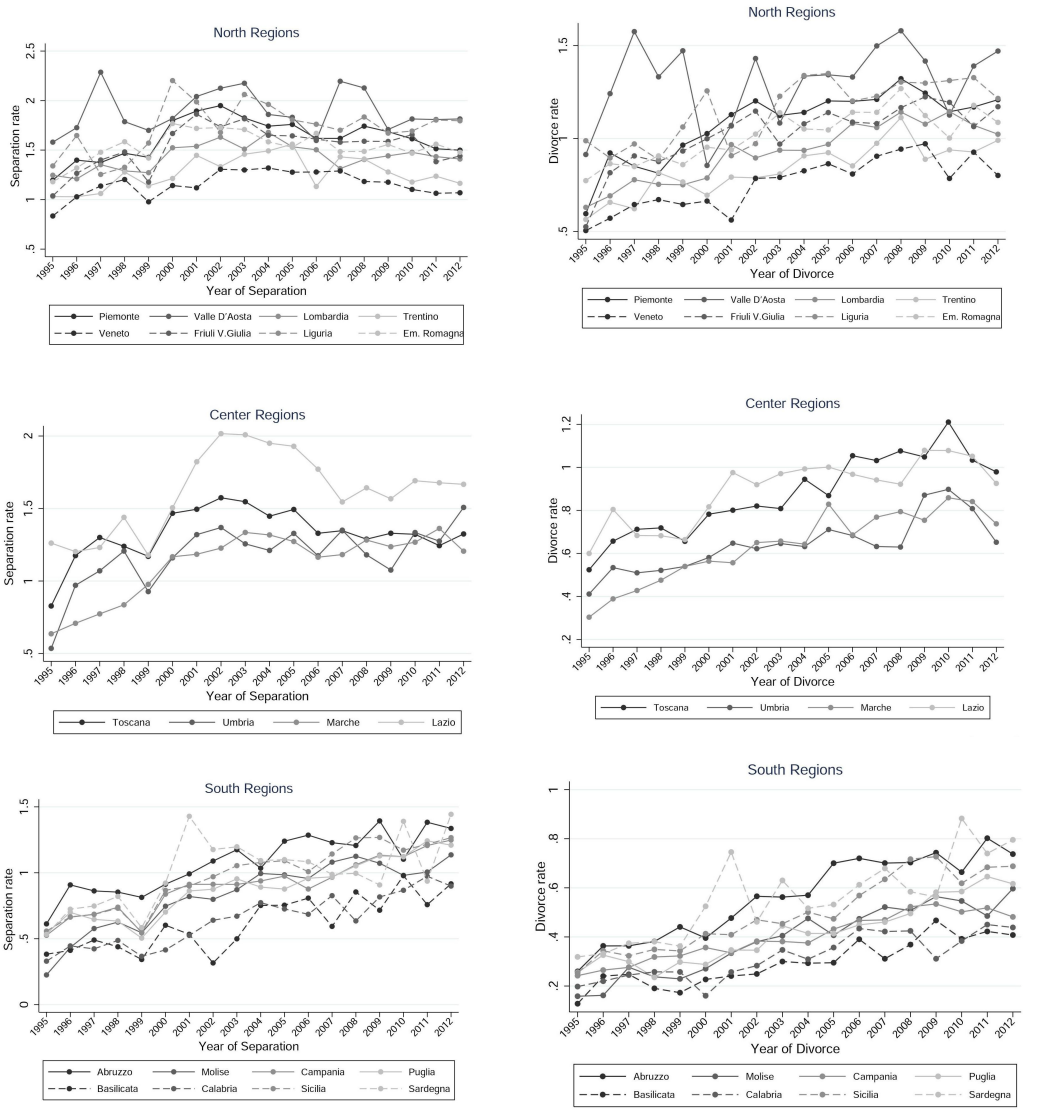
Source: OECD Family Database (1970,1995, 2010), United Nations World Marriage Data and United Nations Demographic Yearbook. The OECD Family Database reports data sources by country. *Figure 1* reports the crude marriage rates (as number of marriages per 1000 people) in OECD countries for the years 1970, 1995 and 2010. Black bars refer to 1970 data, little grey dots report statistics for 1995 and finally dark grey dots refer to 2010. Similarly*Figure 2* reports the crude divorce rates (as number of divorces per 1000 people) in OECD countries for the same years.

Figure A.3: Evolution of Crude Marriage, Separation and Divorce Rates, Italy



Source: ISTAT Data, Demographic Statistics (1970-2012) and ISTAT data from Marriages, Separation and Divorces (1995-2012), Italy. ADELE Elaboration. In Figure are reported the series of the crude marriage, separation and divorce rates from 1970 to 2012 (as absolute number of events per 1000 inhabitants). Marriage rate scale is reported on the left, while Separation and Divorce rate scale is reported on the right.

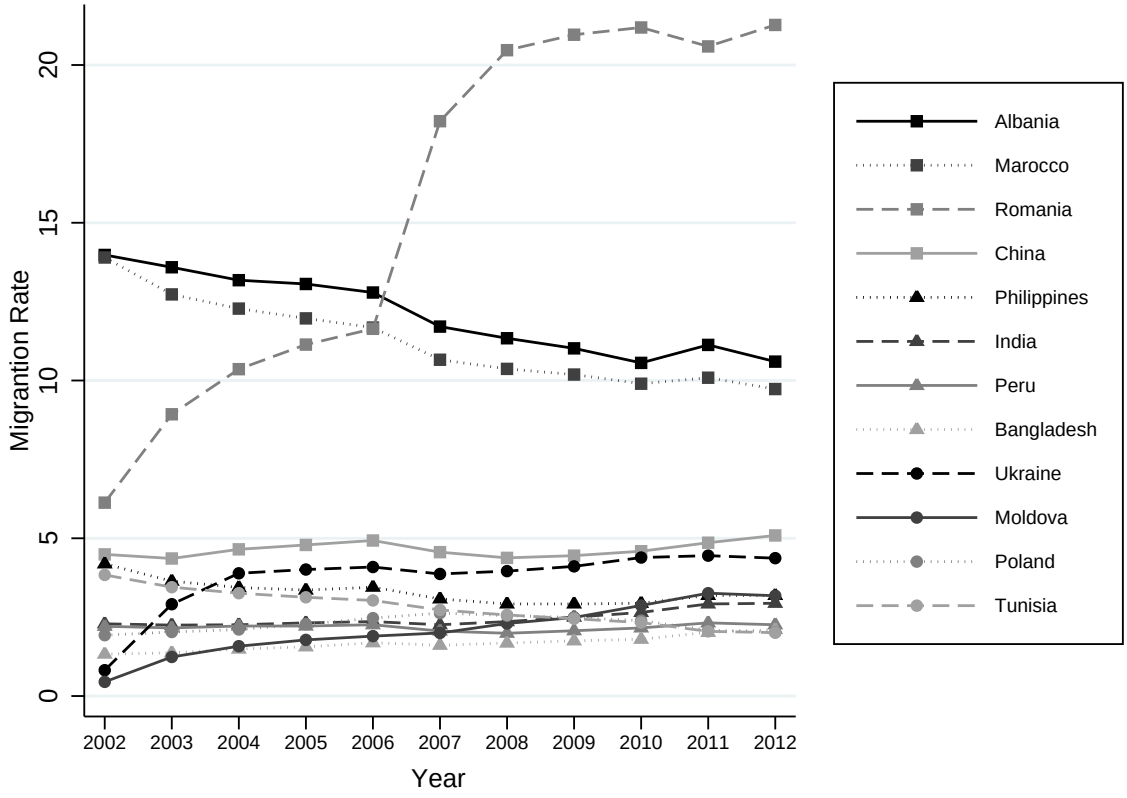
Figure A.4: Evolution of Crude Separation Rates in Italian Regions



Source: Separation and Divorce ISTAT datasets (1995-2012) and Istat Population data - Demographic Report by regions (1995-2012), Italy. In Figure are reported the series of the separation and divorce rates from 1995 to 2012 for Northern, Center and Southern Italian regions (reported on legend). Left panels report separation statistics, right panels report divorce statistics.

A.2 Additional Figures and Tables

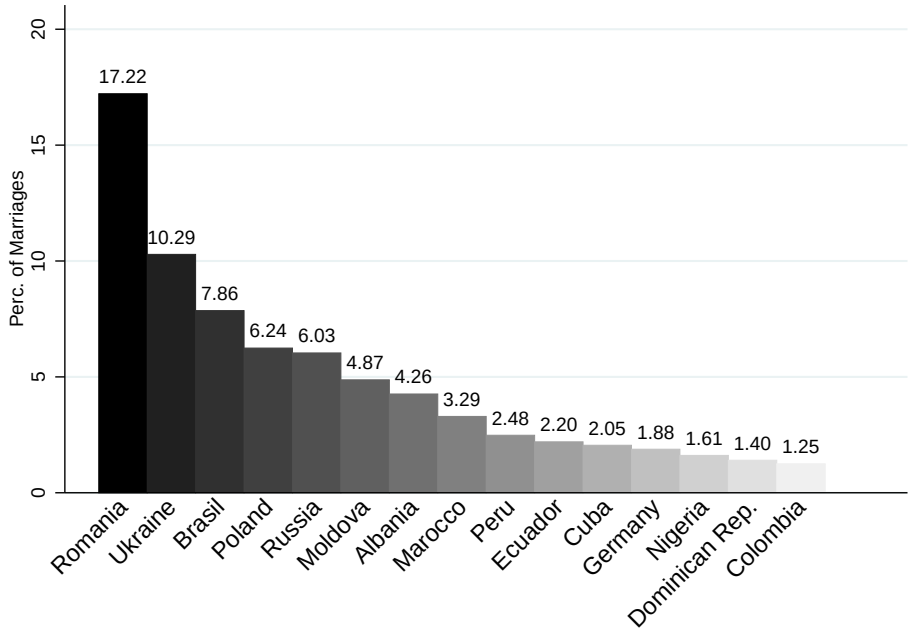
Figure A.5: Evolution of Migratory Inflows by Country of Origin (2002-2012)



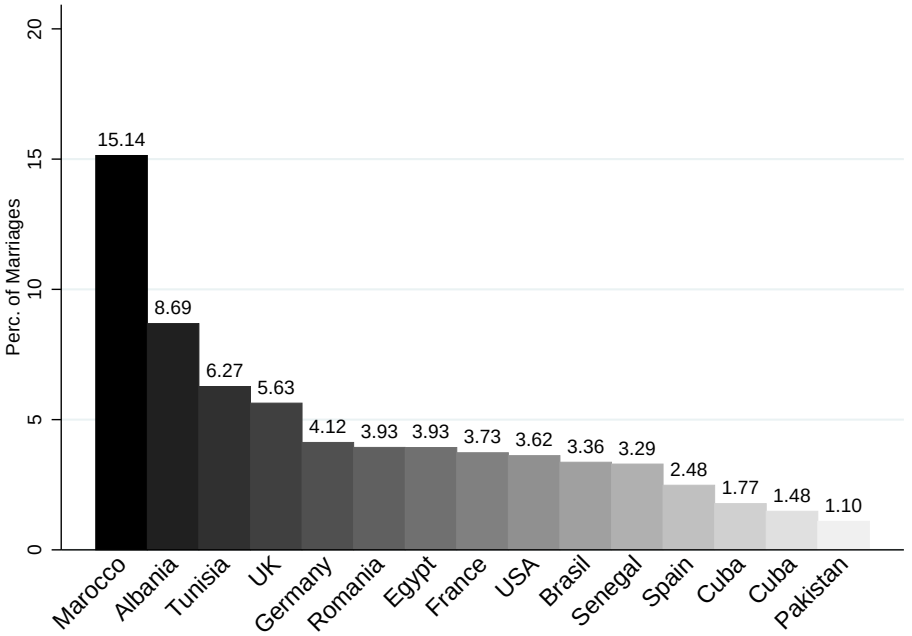
Source: ISTAT - Movement and Annual Evaluation of the Foreign Resident Population by Origin Classification (2002-2012), Italy. Figure show the evolution of the Italian migration inflows by country of origin. The selected countries are the first 12 most represented countries of origin in 2012. The graph reports the share of incoming migrants regularly resident in Italy over the total number of migrant residents by country of origin and by year (2002-2012).

Figure A.6: Heterogamous Marriages by Gender and Country of Origin- Italy (2005-2013)

Panel A: Heterogamous Wife



Panel A: Heterogamous Husband



Source: Marriages. Italy (2005-2013).

Table A.1: Risk of Marital Dissolution: Ethnic Assortative Mating Differences

	Share (%)	Baseline Estimates (1995-2012)			Sub-duration Estimates (1995-2000)	(2001-2006)	(2007-2012)	Citizenship
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic Assortative Mating								
Homogamous(reference group)								
Heterogamous	11.24	.191***	.146***	.136***	.091***	.182***	.163***	.250***
		[1.211]	[1.157]	[1.146]	[1.096]	[1.199]	[1.177]	[1.284]
Migrants	2.01	-.278***	-0.409***	-.375***	-.409***	-.342***	-.411***	-.530***
		[0.757]	[0.664]	[0.687]	[0.665]	[0.710]	[0.663]	[0.589]
Age Assortative Mating								
Wife 16-20 age (refer. group)								
Wife 21-25 age	22.60		-.438***	-.458***	-.478***	-.431***	-.387***	-.455***
			[0.645]	[0.633]	[0.620]	[0.650]	[0.679]	[0.634]
Wife 26-30	37.60		-.782***	-.834***	-.869***	-.787***	-.761***	-.830***
			[0.457]	[0.434]	[0.419]	[0.455]	[0.467]	[0.436]
Wife 31-35	20.70		-1.054***	-1.131***	-1.171***	-1.076***	-1.030***	-1.124***
			[0.349]	[0.323]	[0.310]	[0.341]	[0.357]	[0.325]
Wife 36-40	7.90		-1.284***	-1.369***	-1.003***	-1.286***	-1.220***	-1.361***
			[0.277]	[0.254]	[0.367]	[0.276]	[0.295]	[0.256]
Wife more 40	7.40		-1.744***	-1.816***	-1.793***	-1.703***	-1.575***	-1.806***
			[0.175]	[0.163]	[0.166]	[0.182]	[0.207]	[0.164]
Same age or age diff 0-3 (refer. group)								
Age Diff 4-6	22.30		-.009	-.009	-.023**	.011	.032	-.010*
			[0.991]	[0.991]	[0.977]	[1.010]	[1.033]	[0.990]
Age Diff more 7	19.00		.089***	.080***	.057***	.109***	.120***	.065***
			[1.093]	[1.083]	[1.059]	[1.115]	[1.127]	[1.067]
Age Diff 1-3	11.50		.215***	.220***	.226***	.218***	.203***	.218***
			[1.240]	[1.246]	[1.254]	[1.244]	[1.225]	[1.244]
Age Diff more 4	5.30		.554***	.566***	.547***	.569***	.670***	.555***
			[1.740]	[1.761]	[1.728]	[1.766]	[1.954]	[1.742]
Past Marital History								
Husband First-Wife First (refer. group)								
Husband First - Wife Second	3.50		.489***	.492***	.536***	.457***	.361***	.471***
			[1.631]	[1.636]	[1.709]	[1.579]	[1.435]	[1.602]
Husband Second - Wife First	4.70		0.291***	.275***	.322***	.221***	.218***	.260***
			[1.338]	[1.317]	[1.380]	[1.247]	[1.244]	[1.297]
Husband Second - Wife Second	3.00		.497***	.502***	.565***	.455***	.371***	.487***
			[1.644]	[1.652]	[1.759]	[1.576]	[1.449]	[1.627]

To be Continued...

Table A.2: Risk of Marital Dissolution: Ethnic Assortative Mating Differences

	Share (%)	Baseline Estimates (1995-2012)			Sub-duration Estimates (1995-2000)	(2001-2006)	(2007-2012)	Citizenship
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
...Continued								
Education Assortative Mating								
Husband Low - Wife Low (refer. group)								
Husband Low - Wife Med	15.20		.086***	.059***	.078***	.023*	.028**	.058***
			[1.090]	[1.061]	[1.081]	[1.023]	[1.028]	[1.060]
Husband Low - Wife High	1.90		.063***	.009	.064*	-.045*	-.068	-.005
			[1.065]	[1.009]	[1.066]	[0.956]	[0.934]	[0.995]
Husband Med - Wife Low	6.40		.066***	.039***	.044***	.029*	.024*	.038***
			[1.068]	[1.040]	[1.045]	[1.029]	[1.024]	[1.039]
Husband Med - Wife Med	28.20		.120***	.073***	.109***	.023**	-.040*	.075***
			[1.127]	[1.076]	[1.115]	[1.023]	[0.961]	[1.078]
Husband Med - Wife High	4.90		.063***	-.022*	.031*	-.072***	.150***	-.024*
			[1.065]	[0.978]	[1.031]	[0.931]	[1.162]	[0.976]
Husband High - Wife Low	0.80		.037	-.016	-.056	-.022	.105	-.022
			[1.038]	[0.984]	[1.946]	[0.978]	[1.111]	[0.978]
Husband High - Wife Med	4.60		.052***	-.031**	.023	-.132***	-.039	-.031**
			[1.053]	[0.969]	[1.023]	[0.876]	[1.962]	[0.969]
Husband High - Wife High	7.20		-.037***	-.159***	-.080***	-.235***	-.314***	-.160***
			[0.964]	[0.853]	[0.923]	[0.791]	[0.731]	[0.852]
Disjoint	55.10			.236***	.232***	.226***	.283***	.233***
				[1.266]	[1.261]	[1.254]	[1.327]	[1.262]
Husband Profession	1.673			.013***	.022***	.021***	.030***	.013***
				[1.013]	[1.022]	[1.0211]	[1.030]	[1.013]
Wife Profession	1.39			.038***	.037***	.039***	.027***	.038***
				[1.039]	[1.038]	[1.040]	[1.027]	[1.039]
Observations		4464429	4464429	4464429	1674292	1479605	1310532	4464429
Husband's community pop		No	No	Yes	Yes	Yes	Yes	Yes
Wife's community pop		No	No	Yes	Yes	Yes	Yes	Yes
Region FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (1995-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Family type shares refers to the overall sample. The reference omitted group is of *Homogamous* Italian couples. *Heterogamous* refers to families where one spouse is native and the other one is migrant. *Migrants* is a dummy equal to 1 for families where both spouses are migrant. Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. Columns (4-6) report estimates for specific marriage duration categories. Column (7) exploits citizenship classification of migrants to define family types. Robust standard errors in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Risk of Marital Dissolution: Heterogeneous Effects

	Share (%)	Baseline	Past marital history: First Marriages	Past marital history: Further Marriages	Age categories: Young (h 18-38) (w 15-35)	Age categories: Medium (h 39-49) (w 36-46)	Age categories: Old (h more 50) (w more 47)
		(1)	(2)	(3)	(4)	(5)	(6)
A. Ethnic Assortative Mating							
Homogamous (reference group)							
Heterogamous	11.24	.136***	.134***	.172***	.133***	.135***	.179**
		[1.146]	[1.144]	[1.188]	[1.143]	[1.144]	[1.196]
		(.006)	(.006)	(.013)	(.007)	(.020)	(.069)
Migrants	2.01	-.375***	-.393***	-.281***	-.509***	-.005	.283
		[0.687]	[0.675]	[0.755]	[0.601]	[0.995]	[1.328]
		(.016)	(.018)	(.039)	(.020)	(.055)	(.20)
B. Ethnic Assortative Mating: Gender gap							
Homogamous (reference group)							
Heterogamous Wife	7.79	.117***	.127***	.128***	.140***	.120***	.201**
		[1.125]	[1.135]	[1.137]	[1.151]	[1.127]	[1.223]
		(.006)	(.006)	(.013)	(.007)	(.020)	(.069)
Heterogamous Husband	3.45	.173***	.147***	.334***	.122***	.183***	.035
		[1.189]	[1.158]	[1.396]	[1.130]	[1.200]	[1.036]
		(.009)	(.010)	(.024)	(.011)	(.038)	(.181)
Observations		4464429	3968860	495569	3172286	487130	76815
Assortative mating controls		Yes	Yes	Yes	Yes	Yes	Yes
Labour market controls		Yes	Yes	Yes	Yes	Yes	Yes
Husband's community population		Yes	Yes	Yes	Yes	Yes	Yes
Wife's community population		Yes	Yes	Yes	Yes	Yes	Yes
Region FE			Yes	Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (1995-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Family type shares refers to the overall sample. In Panel A: the reference omitted group is of *Homogamous* Italian families; *Heterogamous* refers to families where one spouse is native and the other one is migrant; *Migrants* is a dummy equal to 1 for families where both spouses are migrant. In Panel B: the reference omitted group is of *Homogamous* Italian families; *Heterogamous Wife* are families where the wife is born abroad Italy, reversely *Heterogamous Husband* are families where the husband is the foreign spouse; *Migrants* estimates are not reported in the bottom Panel. Column (1) reports baseline estimates as in Table 1.2. Column (2) reports estimates for first marriages only while Column (3) reports estimates for second or further marriages of one or both spouses. Columns (4-6) report estimates for specified age categories. Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. Columns (4-6) report estimates for specific marriage duration categories. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.4: Countries Classification by Macroarea

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

Note: ISTAT Classification of foreign countries by macroarea 2012. Minor changes intervene to modify the classification over time. The changes include the establishment of a new state (e.g. Indonesia or Sudan); the termination of a pre-existing state (e.g. Serbia and Montenegro because of Montenegro separation); the change of geopolitical area of membership status (as in the case of the States that entered over time into EU); the introduction and/or the exclusion of a new entry in adaptation to international guidelines. No change modifies significantly the Classification adopted in the analysis. For a detailed description of changes of Foreign Countries Classification starting from 2002 refer to http://www.istat.it/it/archivio/6747.

Table A.5: Correlation Matrix of Cultural Distance Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Genetic Distance I (Fst)	1									
(2) Genetic Distance II (Nei)	.992	1								
(3) Genetic Distance I (Fst weight)	.906	.898	1							
(4) Genetic Distance II (Nei weight)	.898	.905	.991	1						
(5) Linguistic Distance I (Ling. Tree)	.617	.581	.538	.505	1					
(6) Linguistic Distance II (ASJP)	.573	.544	.481	.455	.983	1				
(7) Geo-Linguistic Distance I (dist w)	.751	.744	.652	.644	.685	.665	1			
(8) Geo-Linguistic Distance II (dist w)	.713	.711	.593	.588	.658	.668	.979	1		
(9) Geo-Linguistic Distance I (dist ces)	.752	.745	.654	.646	.674	.654	.997	.978	1	
(10) Geo-Linguistic Distance II (dist ces)	.715	.713	.595	.591	.647	.655	.979	.999	.978	1

Note: Marriage and Separations ISTAT datasets, 1995-2012, Italy. ADELE elaboration. Matrix of correlations among Cultural Distance proxies. Cultural distance is measured in terms of genetic and ethnolinguistic distance. The genetic distance metrics (*Genetic distance I* and *Genetic distance II*) are constructed on the *coancestor coefficients*: indices of heterozygosity, i.e. the probability that two alleles at a given locus selected at random from two populations will be different (Cavalli-Sforza and Piazza, 1994) (*Source*: Spolaore and Wacziarg (2009)). The ethnolinguistic distance measures are based on the language tree classification and on lexicostatistics analysis. The first ethnolinguistic variable (*Linguistic distance I*) is based on the linguistic tree categorization proposed by Lewis et al. (2009). The second ethnolinguistic variable (*Linguistic distance II*) derive from the Automated Similarity Judgment Program (ASJP) (*Source*: Melitz and Toubal (2014b) and Egger and Toubal (2016)). To account for geographical relatedness and potential interactions, new weighted ethnolinguistic variables have been constructed: *Geo-linguistic distance I* and *Geo-linguistic distance II* where the ethnolinguistic distances are weighted with respect to the relative geographical distance between the countries [weighted and CES geographical distances exploited]. For a more accurate description of the variables construction refer to the above mentioned references.

Table A.6: Risk of Marital Dissolution: Cultural Distance Proxies

	Genetic Distance I (*Fst* Definition) (1)	Genetic Distance II (*Nei* Definition) (2)	Linguistic Distance I (Ling. Tree) (3)	Linguistic Distance II (ASJP) (4)	Geo-Linguistic Distance I (Weight) (5)	Geo-Linguistic Distance II (CES Weight) (6)
Mean	40.08	7.34	.43	.66	.078	.13
	(176.33)	(34.28)	(1.26)	(1.86)	(.33)	(.48)
Estimates	.143***	.123***	.087***	.099***	.177***	.179***
	[1.154]	[1.131]	[1.090]	[1.104]	[1.194]	[1.194]
	(.010)	(.010)	(.004)	(.004)	(.009)	(.012)
$Estimates^2$	-.013***	-.010***	-.010*	-.011	-.020***	-.023***
	[0.987]	[0.990]	[0.990]	[0.987]	[0.987]	[0.977]
	(.002)	(.001)	(.014)	(.008)	(.002)	(.002)
Observations	1743834	1743834	1743834	1743834	1743834	1743834
Assortative mating controls	Yes	Yes	Yes	Yes	Yes	Yes
Labour market controls	Yes	Yes	Yes	Yes	Yes	Yes
Husband's community population	Yes	Yes	Yes	Yes	Yes	Yes
Wife's community population	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Marriage and Separations ISTAT datasets (2005-2012), Italy. ADELE elaboration. Cox regression coefficients estimates reported. Hazard rates reported in squared brackets. Each column corresponds to a different estimation. Cultural distance is measured in terms of genetic and ethnolinguistic distance. The genetic distance metrics (*Genetic distance I* and *Genetic distance II*) are constructed on the *coancestor coefficients*: indices of heterozygosity, i.e. the probability that two alleles at a given locus selected at random from two populations will be different (Cavalli-Sforza and Piazza, 1994) (*Source*: Spolaore and Wacziarg (2009)). The ethnolinguistic distance measures are based on the language tree classification and on lexicostatistics analysis. The first ethnolinguistic variable (*Linguistic distance I*) is based on the linguistic tree categorization proposed by Lewis et al. (2009). The second ethnolinguistic variable (*Linguistic distance II*) derive from the Automated Similarity Judgment Program (ASJP) (*Source*: Melitz and Toubal (2014b) and Egger and Toubal (2016)). To account for geographical relatedness and potential interactions, new weighted ethnolinguistic variables have been constructed: *Geo-linguistic distance I* and *Geo-linguistic distance II* where the ethnolinguistic distances are weighted with respect to the relative geographical distance between the countries. For a more accurate description of the variables construction refer to the above mentioned references. All variables are standardized. Table reports mean and standard deviation of origin variables. Region and year of marriage fixed effects are included in all specifications. *Assortative mating controls* account for the most relevant assortative mating dimensions in terms of age, education and past marital history of spouses, variables are described in Table 1.1. *Labour market controls* include professional status (blue-collar, white-collar, director, self employed and entrepreneur) of both spouses. The specification adds to marriage controls a dummy for prenuptial financial agreement (disjoint vs joint management of family wealth). *Husband and Wife community pop* control for the population size of the spouses province of residence as a proxy of the local marriage market of reference, at the time of marriage. Robust standard errors are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Appendix B

A Study of Marriage, Fertility and Divorce: Cultural-Ethnic Socialization of Migrants in Italy

B.1 Theoretical Model: Derivations

B.1.1 The Primal and The Dual Matching Problem

In TU models, there is a correspondence between the equilibrium concept of stability and the solution of a welfare maximization. For instance, whenever a stable matching exists, it maximizes the total marital output over all potential assignments in the market [Shapley and Shubik (1971)]. The welfare maximization problem could be stated as:

$$\max_{\mu(\{i,j\})\geq 0} \quad \sum_{ij} \mu\left(\{i,j\}\right) \tilde{U}\left(\{i,j\}, q^i\right) + \epsilon_m^{ij} + \eta_f^{ij} \tag{B.1}$$

subject to:

$$\begin{aligned} \sum_j \mu\left(\{i,j\}, q^i\right) + \mu\left(\{i,0\}, q^i\right) &= m_i \quad \forall i \in \{a,b\}, \\ \sum_i \mu\left(\{i,j\}, q^i\right) + \mu\left(\{0,j\}, q^i\right) &= f_j \quad \forall j \in \{a,b\}, \\ \mu\left(\{i,j\}, q^i\right) \geq 0, \mu\left(\{i,0\}, q^i\right) \geq 0, \mu\left(\{0,j\}, q^i\right) &\geq 0. \end{aligned} \tag{B.2}$$

Following Galichon and Salanié (2015), we could also restate the problem in (B.1), underlining the difference between the systematic marital component and the individual stochastic dimensions:

$$\max_{\mu(\{i,j\})\geq 0} \quad \sum_{ij} \tilde{U}\left(\{i,j\},q^i\right) - \varepsilon(\mu) \tag{B.3}$$

where:

$$\varepsilon(\mu) = \alpha^*(\mu) + \beta^*(\mu)$$

$$\alpha^*(\mu) = \max_{U_m(\{ij\})} \left\{ \sum_{ij} \mu\left(\{i,j\}\right) \tilde{U}_m\left(\{i,j\},q^i\right) - \alpha(U_m^i) \right\}, \quad \alpha(U_m^i) = \sum_{i\in\{a,b\}} H(U_m)$$

and

$$\beta^*(\mu) = \max_{U_f(\{ij\})} \left\{ \sum_{ij} \mu\left(\{i,j\}\right) \tilde{U}_f\left(\{i,j\},q^i\right) - \beta(U_f^j) \right\}, \quad \beta(U_f^j) = \sum_{j\in\{a,b\}} Q(U_f).$$

$\alpha^*(\mu)$ represents the Legendre transform of $\alpha(U_m^i)$ and similarly for $\beta^*(\mu)$. In equation (B.1), the first term reflects the aggregate systematic preferences. The second term, instead, reflects the dispersion of individual preferences with respect to observable characteristics. For instance, $\alpha^*(\mu)$ and $\beta^*(\mu)$ correspond to the amount of heterogeneity we need to introduce in the model to rationalize observed individual choices. In our specific logit formulation, those components characterize the entropy of the discrete choice problems. For an extended presentation of the problem see Galichon and Salanié (2015).

The welfare maximization problem is linear in the optimal assignment, hence it admits a dual. By duality results, if a solution to the welfare maximization problem exists, it could be achieved by the decentralized equilibrium implied by the related linear dual problem. Under its dual form the optimal marital matching is characterized by a cost minimization problem. It identifies the set of possible divisions of marital value between spouses. The dual problem formulates as:

$$\min_{U_m^i,U_f^j} \quad \sum_{i\in\{a,b\}} m^i U_m^i + \sum_{j\in\{a,b\}} f^j U_f^j \tag{B.4}$$

subject to:

$$\tilde{U}_m\left(\{i,j\},q^i\right) + \tilde{U}_f\left(\{i,j\},q^i\right) \geq \tilde{U}\left(\{i,j\},q^i\right) \quad \left[\mu\left(\{i,j\},q^i\right)\right] \tag{B.5}$$

We can easily interpret the constraint in (B.5) as a stability condition. Namely, whenever the constraint hold with an equality, it implies that the value generated from the marital union in equivalent to the sum of spouses utilities, hence a matched is formed. On the opposite case when the constraint is not binding it implies that individuals might be better

off under a different assignment, which contradicts the stability condition (because of the presence of blocking-pairs). The Lagrange multiplier associated to the stability constraints its exactly the assignment probability $\mu\left(\{i,j\},q^i\right)$. Again, by complementary slackness, $\mu\left(\{i,j\},q^i\right)>0$, when the sum of spouses indirect utilities correspond to the actual break-down of the marital surplus; while $\mu\left(\{i,j\},q^i\right)=0$ when the constraint is slack.

B.1.2 Multinomial Logit Model Derivation

Considering a multidimensional framework of K cultural traits in the society ($K>2$: $i,j,k=1,..,K$), the discrete choice problem of man m with trait i is specified by:

$$U_m^i=\tilde{U}_m\left(\{i,j\},q^i\right)+\epsilon_m^{ij},\quad \forall i,j,k=1,..,K$$

The chosen alternative j is the one that leads the larger utility in such a way that:

$$\begin{aligned} Pr[j] &= Pr\left[\tilde{U}_m\left(\{i,j\},q^i\right)+\epsilon_m^{ij}\geq \tilde{U}_m\left(\{i,k\},q^i\right)+\epsilon_m^{ik},\quad \forall k\neq j\right] && \text{(B.6)}\\ &= Pr\left[\epsilon_m^{ik}\leq \tilde{U}_m\left(\{i,j\},q^i\right)-\tilde{U}_m\left(\{i,k\},q^i\right)+\epsilon_m^{ij},\quad \forall k\neq j\right] && \text{(B.7)}\end{aligned}$$

We assume that ϵ_m^{ij} is an independent and identically distributed random variable with extreme-value cumulative distribution, such as $G(\epsilon)=\exp(-\exp(-\epsilon))$ and density equal to: $g(\epsilon)=\exp(-\epsilon\exp(-\epsilon))$. Hence the probability that the woman with trait j is chosen by man m with cultural trait i might be written as:

$$\begin{aligned} Pr[j] &= \int\prod_{k\neq j}G\left(\tilde{U}_m\left(\{i,j\},q^i\right)-\tilde{U}_m\left(\{i,k\},q^i\right)+\epsilon_m^{ij}\right)g(\epsilon_m^{ij})d\epsilon_m^{ij}\\ &= \int\prod_{k\neq j}\exp\left(-\exp\left(-\left(\tilde{U}_m\left(\{i,j\},q^i\right)-\tilde{U}_m\left(\{i,k\},q^i\right)+\epsilon_m^{ij}\right)\right)\right)\exp(-\epsilon_m^{ij}-\exp(-\epsilon_m^{ij}))d\epsilon_m^{ij}\\ &= \int\exp(-\epsilon_m^{ij})\prod_{k\neq j}\exp\left(-\exp\left(-\left(\tilde{U}_m\left(\{i,j\},q^i\right)-\tilde{U}_m\left(\{i,k\},q^i\right)+\epsilon_m^{ij}\right)\right)\right)\exp(-\exp(-\epsilon_m^{ij}))d\epsilon_m^{ij}\\ &= \int\exp(-\epsilon_m^{ij})\exp\left(\sum_{k\neq j}-\exp\left(-\left(\tilde{U}_m\left(\{i,j\},q^i\right)-\tilde{U}_m\left(\{i,k\},q^i\right)+\epsilon_m^{ij}\right)\right)\right)\exp(-\exp(-\epsilon_m^{ij}))d\epsilon_m^{ij}\\ &= \int\exp(-\epsilon_m^{ij})\left\{\exp\left(-\exp\left(-(\epsilon_m^{ij})\left(1+\sum_{k\neq j}\exp(-\left(\tilde{U}_m\left(\{i,j\},q^i\right)-\tilde{U}_m\left(\{i,k\},q^i\right)\right)\right)\right)\right)\right\}d\epsilon_m^{ij}\end{aligned}$$

Given that $\int\alpha\exp(-\epsilon)\exp(-\alpha\exp(-\epsilon))d\epsilon=1$, the probability that the woman with trait j is chosen by man m with cultural trait i translates into:

$$Pr[j]=\frac{\exp\left(\tilde{U}_m\left(\{i,j\},q^i\right)\right)}{1+\sum_k\exp\left(\tilde{U}_m\left(\{i,k\},q^i\right)\right)}$$

B.1.3 Proofs of Propositions

Proof of Proposition 1. Under the assumption that $c(\tau)$ is a monotonic increasing and convex function in its argument, the optimization problem in (2.3) is a convex problem, which guarantees that optimal socialization probabilities are continuous functions in the parameters, $\tau\left(\{i,j\},d,q^i\right)$. The implicit function theorem on the first order condition of problem (2.3) implies that $\partial\tau\left(\{i,j\},d,q^i\right)/\partial q^i < 0$.

Proof of Proposition 2 and 3. Let $\pi\left(\{i,j\},n,q^i\right) = \pi\left(\bar{\theta}\right)$, where

$$\bar{\theta} \equiv (n\tilde{W}(\{i,j\},1,q^i) - n\tilde{W}(\{i,j\},0,q^i)).$$

$\pi\left(\bar{\theta}\right)$ is a monotonic increasing function in $\bar{\theta}$, being the truncated cumulative distribution of θ up to the point $\bar{\theta}$. Monotonicity implies that if a couple remains married for some realization $\hat{\theta}$, then it also do so for any θ, such that $\theta > \hat{\theta}$. Results of Proposition 2 and 3 follow from straightforward calculation of $\bar{\theta}\left(\{i,j\},n,q^i \mid i=j\right)$ and $\bar{\theta}\left(\{i,j\},n,q^i \mid i\neq j\right)$ from marital indirect utilities derived from the socialization process.

Proof of Proposition 4. Let rewrite the optimization problem in (2.4), as

$$\max_n U\left(\left(\{i,j\},n,q^i\right)\right) = \left[\begin{array}{l} \pi\left(\{i,j\},n,q^i\right) n\tilde{W}(\{i,j\},1,q^i) + \\ \left(1-\pi\left(\{i,j\},n,q^i\right)\right)\left(n\tilde{W}(\{i,j\},0,q^i)+\theta\right) - c(n) \end{array}\right]$$

$$\max_n U\left(\left(\{i,j\},n,q^i\right)\right) = \left[\begin{array}{l} \int_{\bar{\theta}}^{\infty} n\tilde{W}(\{i,j\},1,q^i) f(\theta)d\theta + \\ \left(1-\pi\left(\{i,j\},n,q^i\right)\right)\left(n\tilde{W}(\{i,j\},0,q^i)+\theta\right) - c(n) \end{array}\right]$$

For quadratic fertility costs, the first order condition wrt n is:

$$foc_n:\ n = \left[\begin{array}{l} f(\bar{\theta})\left\{n\tilde{W}(\{i,j\},1,q^i) - n\tilde{W}(\{i,j\},0,q^i) - \bar{\theta}\right\} + \tilde{W}(\{i,j\},0,q^i) + \\ \pi\left(\{i,j\},n,q^i\right)\left\{\tilde{W}(\{i,j\},1,q^i) - \tilde{W}(\{i,j\},0,q^i)\right\} \end{array}\right],$$

as the derivative of an integral wrt the lower bound equals the value of the integrand at that point. Cancelling out terms, we obtain:

$$foc_n:\ n = \tilde{W}(\{i,j\},0,q^i) + \pi\left(\{i,j\},n,q^i\right)\left\{\tilde{W}(\{i,j\},1,q^i) - \tilde{W}(\{i,j\},0,q^i)\right\}.$$

B.2 Data and Sample Definition

The empirical analysis uses administrative Italian data at the individual level, from 1995 to 2012. We provide, in what follow, a brief description of the data sources and the main information provided in each dataset.

Marriage distribution. We exploit marriage records from municipality 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 celebration ceremony (religious or civil), the municipality of celebration and the choice of the property regime by the spouses (community or separation property). The information provided for each spouse include: 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 migrant 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.

Unmatched Individuals. We derive the population vectors by ethnic group, sex 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 case 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 as described in Section (3.5).

Fertility rate. 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 related to the born as sex, date and province of birth, citizenship and parents information regarding their date of birth, citizenship and marital status.

Separation rate. 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. The statistical data collected allow to analyse different aspects of the marital dissolution phenomenon. Information are provided regarding the judicial and appeal to legal assistance proceedings; the marital union as the date and the celebration ceremony; spouses

socio-demographic characteristics as those reported in marriage records; any children involved as date of birth and sex; the post-dissolution arrangements as alimony obligations, recipient subject and annual amount of contribution to the maintenance, custody assignment of children [1]. We focus on separation rates, which better represent marital dissolution decisions in the Italian context, for two main reasons. First, on average only the 65% of separations are followed by a divorce, which implies that divorce choices underestimate marital dissolution behaviours. Secondly, Law 74/1987 fixed a minimum period of three years before a divorce submission after separation. This implies that marital duration is overestimated considering divorce and above all, because of the censored structure of our data, marital dissolution decisions might be underestimated.

Data construction and Sample definition: The empirical estimation is based on a unique quasi-longitudinal dataset that links households information across different sources. Time invariant dimensions have been exploited to associate marriages with birth and separation records: the exact date of the marriage and the exact date and province of birth of both spouses. For the matching procedure, we account for changes in the classification of provinces over time. In the birth records matching, the combination of these characteristics allows to drop only the 1.2% of marriages, while in the separation matching, we drop the 0.5% of marriages. Such low percentages suggest that marriages can be uniquely identified through the set of time-invariant characteristics listed above [2]. The total sample of marriages (4,151,528) correspond to the 92.58% of the universe of marriages celebrated in Italy during the time interval 1995-2012.

From this final sample, we recover the following empirical moments. Marital gains $\hat{\varphi}_g(\{i,j\})$ are computed from equation (2.7), exploiting the number of i,j marriages formed in each province g ($\mu\{i,j\}$), and the number of available individuals i and j for each province g, ($\mu\{i,0\}, \mu\{0,j\}$). Fertility rates $\hat{n}_g(\{i,j\})$ for each $\{i,j\}$ match and for all provinces g are computer as:

$$\hat{n}_g(\{i,j\}) = \frac{1}{M_g(\{i,j\})} \sum_{m=1}^{M_g(\{i,j\})} N_g(\{i,j\})$$

[1] For the period under investigation, registries of civil court chancelleries constitute the unique source for separations and divorces data, while starting from December 2014 (in application of Law n. 162/2014) consensual separation and divorce proceedings can be submitted to the civic registrar. The time-period of our analysis rules out potential concerns in the selection of available data. Proceedings classified to end with conciliation, cancellation, or change of rite are registered, but post-dissolution information is not available for them. We drop them from the final sample because not representative of effective households' marital dissolution choices. Separation records that end up in conciliation are 2,149, those cancelled are 1,884 and those that changed rite are 1,772; hence they account of the 1.59% of the total number of separations.

[2] The exact matching procedure might still suffer from coding errors. Moreover separations related to marriages celebrated abroad cannot be traced in vital records registries and hence excluded from the sample. Among the not-merged separations, the number of heterogamous families is overrepresented, as expected, because those couples are more likely to have celebrated their marriage abroad.

where $N_g(\{i,j\})$ represents the number of children born from an $\{i,j\}$ marriage and $M_g(\{i,j\})$ is the total number of $\{i,j\}$ marriages formed in province g.

Separations rates $\hat{\pi}_g(\{i,j\})$ for each $\{i,j\}$ match and for all provinces are computed as:

$$\hat{\pi}_g(\{i,j\}) = \frac{1}{M_g(\{i,j\})} \sum_{b=1}^{M_g(\{i,j\})} D_g(\{i,j\})$$

where $D_g(\{i,j\})$ is a dummy equal to one if the $\{i,j\}$ marriage dissolved and $M_g(\{i,j\})$ is the total number of $\{i,j\}$ marriages formed in province g.

Socialization probabilities: Socialization data come from the *Condition and Social Integration of Foreign Nationals* Survey, conducted in 2011 and 2012 in all Italian provinces on a sample of 9600 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 relevant aspects of the socio-economic integration process of migrants in Italy, with a particular focus on the linguistic integration. Different dimensions have been targeted such as: the family composition, the educational level, the migratory path, the employment status, the discrimination and integration perception, the living environment conditions, the religious affiliation, the social network formation and the 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, that requires also the participation of sociologists and cultural mediators. The final questionnaire was translated in 10 different languages to overcome potential language barriers and reduce the attrition rate. The actual survey has been conducted through direct interviews supported by the CAPI (Computer Assisted Personal Interview) system. In each selected household, all members have been interviewed, either foreign-born and natives. We exclude from our analysis, respondents who are not married and families without children, at the time of the interview. The final sample, we exploit to derive socialization frequencies, is of about 17.512 individuals belonging to 4996 families, and the 18.59% of those families are either separated or divorced. We consider the sample as a representative sample to study migrants linguistic integration by ethnic group in each province of residence. We construct our measure of socialization based on the language spoken at home [3]. The survey also provide questions to evaluate the level of Italian language proficiency and we check individual self-declared responses on language spoken.

Unfortunately, the sample of families ended in separation is not sufficient to precisely estimate socialization frequencies by family type and province. Hence we compute the vector of socialization frequencies $\hat{P}_{k,g}(\{i,j\},d)$ for all i,j and k, for marital status d and for

[3] Examples of the questionnaire and invitation letter are available from *http://www.istat.it/it/archivio/10825*.

all provinces g, as follow:

$$\hat{P}_{k,g}(\{i,j\}, d=0) = \frac{1}{M_g(\{i,j\})} \sum_{b=1}^{M_g(\{i,j\})} S_{k,g}(\{i,j\})$$

$$\hat{P}_{k,g}(\{i,j\}, d=1) = \widehat{\alpha^{ij}} \hat{P}_{k,g}(\{i,j\}, d=0) \quad \forall g$$

where: $\widehat{\alpha^{ij}}$ is computed as the average difference between the observed socialization probabilities in married as compared to divorced families, for each i, j and k.

Population distribution: Finally, the population distribution by ethnic group for each province is derived from municipality records on the movements of the foreign resident population for years from 1995 to 2010. Ethnic group shares are calculated thanks to municipality data on the total resident population, aggregated at the province level.

Table B.1: Data Description: sources and variables of interest

Data	Time-Frequency	Source	Variables of Interest
Marriage records	Year (1995-2012)	Municipality Vital Statistics Registries	Wedding: the date of marriage, the celebration ceremony (religious or civil), the municipality of celebration, the property regime (community or separation property). Spouses: date of birth, municipality of birth, municipality of residence at the time of marriage, municipality of future residence, past marital status, the education level, the employment status, the nationality, citizenship acquisition (Italian born, naturalized Italian and not Italian), the country of origin if foreign
Birth records	Year (1990-2012)	Municipality Births Registries	Born: sex, date and municipality of birth, citizenship, family size, presence and number of minor members in the family. Parents: date and province of birth, citizenship and marital status
Separation records	Year (1995-2012)	Registries of Civil Court Chancelleries	Proceeding: judicial and appeal to legal assistance, proceeding end, date of registration, date of separation, court of reference; Marriage: date of marriage, celebration ceremony (religious or civil); Spouses: date of birth, municipality of birth, municipality of residence at the time of marriage, the municipality of future residence, past marital status, the education level, the employment status, the nationality, citizenship acquisition (Italian born, naturalized Italian and not Italian), the country of origin if foreign; Children: number of children born in the marriage, number of children in the family at separation, date of birth and sex; Post-dissolution arrangements (2000-2012): alimony obligations versus children and/or spouse (yes or no), recipient subject and annual amount of contribution to the maintenance, custody assignment of children.
Socialization	(2011-2012)	Survey: Condition and Social Integration of Foreign Nationals	Household Panel. Individual data: age, sex, relationship with targeted subject, marital status, year of marriage, nationality, citizenship acquisition (Italian born, naturalized Italian and not Italian), country of origin if foreign, partner/mother/father citizenship and country of origin, migratory path, educational level, employment status/ school enrolment, religious affiliation, discrimination and integration perception, social networks (at work, school, freetime). Household data: family composition, area of residence (province), living environment conditions. Language data: first language, verbal and written knowledge of first language, language spoken at home, Italian language proficiency: lecture, writing, reading, comprehension level (good, enough, little, nothing)
Migration records	Year (1995-2010)	Municipality Population Balance	End of period data (December, 31): total population, total foreign population, total male and female foreign population by municipality. Male and female foreign population by country of origin by province.

B.3 The Structural Model

Socialization probabilities in:
Italian homogamous families $\{I, I\}$:

$$P_{I,g}(\tau; \{I, I\}, 0, Q_g^I) = 1$$

Minorities homogamous families $\{i, j\}$, $i = j$:

$$P_{i,g}(\tau; \{i, j\}, 0, Q_g^i) = \tilde{\tau}_{m,g}^i + \tilde{\tau}_{f,g}^j(1 - \tilde{\tau}_{m,g}^i - \tilde{\tau}_{f,g}^j)Q_g^i \qquad \forall i = j \quad i \neq I$$

$$P_{I,g}(\tau; \{i, j\}, 0, Q_g^i) = (1 - \tilde{\tau}_{m,g}^i - \tilde{\tau}_{f,g}^j)Q_g^I \qquad \forall i = j \quad i \neq I$$

$$\frac{\partial C(\tau_g)}{\partial \tau_g} = 2\Delta V^{iI} Q_g^I \quad \forall i$$

Minorities heterogamous families $\{i, I\}$:

$$P_{i,g}(\tau; \{i, I\}, 0, Q_g^i) = \tilde{\tau}_{m,g}^i + (1 - \tilde{\tau}_{m,g}^i - \tilde{\tau}_{f,g}^I)Q_g^i \qquad \forall i \neq I$$

$$P_{I,g}(\tau; \{i, I\}, 0, Q_g^i) = \tilde{\tau}_{f,g}^I + (1 - \tilde{\tau}_{m,g}^i - \tilde{\tau}_{f,g}^I)Q_g^I \qquad \forall i \neq I$$

$$\tau_{m,g}^i > 0; \quad \frac{\partial C(\tau_g)}{\partial \tau_g} = Q_g^I(\Delta V^{iI} - \Delta V^{Ii})$$

Minorities heterogamous families $\{I, j\}$:

$$P_{I,g}(\tau; \{I, j\}, 0, Q_g^I) = \tilde{\tau}_{m,g}^I + (1 - \tilde{\tau}_{m,g}^I - \tilde{\tau}_{f,g}^j)Q_g^I \qquad \forall j \neq I$$

$$P_{j,g}(\tau; \{I, j\}, 0, Q_g^j) = \tilde{\tau}_{f,g}^j + (1 - \tilde{\tau}_{m,g}^I - \tilde{\tau}_{f,g}^j)Q_g^j \qquad \forall j \neq I$$

$$\tau_{f,g}^j > 0; \quad \frac{\partial C(\tau_g)}{\partial \tau_g} = Q_g^I(\Delta V^{jI} - \Delta V^{Ij})$$

Minorities heterogamous families $\{i, j\}$, $i \neq j$:

$$P_{i,g}(\tau; \{i, j\}, 0, Q_g^i) = \tau_{m,g}^i + (1 - \tau_{m,g}^i - \tau_{f,g}^j)Q_g^i \qquad \forall i, j \neq I$$

$$P_{j,g}(\tau; \{i, j\}, 0, Q_g^j) = \tau_{f,g}^j + (1 - \tau_{m,g}^i - \tau_{f,g}^j)Q_g^j \qquad \forall i, j \neq I$$

$$P_{I,g}(\tau; \{i, j\}, 0, Q_g^I) = (1 - \tau_{m,g}^i - \tau_{f,g}^j)Q_g^I \qquad \forall i, j \neq I$$

$$\tau_{f,g}^j > 0; \quad \frac{\partial C(\tau_{f,g}^j)}{\partial \tau_{f,g}^j} = Q_g^i(\Delta V^{ji} - \Delta V^{ij}) + Q_g^I(\Delta V^{jI} + V_j^i - V_I^i)$$

$$\tau^i_{m,g} > 0; \quad \frac{\partial C(\tau^i_{m,g})}{\partial \tau^i_{m,g}} = Q^j_g(\Delta V^{ij} - \Delta V^{ji}) + Q^I_g(\Delta V^{iI} + V^j_i - V^j_I)$$

Divorce probabilities:

$$\pi_g\left(\{i,j\}, n, Q^i_g\right) = \frac{\exp \bar{\theta}_g(\{i,j\})}{1 + \exp \bar{\theta}_g(\{i,j\})},$$

with

$$\bar{\theta}_g(\{i,j\}) \equiv (n_g \tilde{W}_g(\{i,j\}, 1, Q^i_g) - n_g \tilde{W}_g(\{i,j\}, 0, Q^i_g)) \quad \forall i,j.$$

Fertility rates:

$$n_g = \tilde{W}_g(\{i,j\}, 0, Q^i_g) + \pi_g\left(\{i,j\}, n_g, Q^i_g\right)\left\{\tilde{W}_g(\{i,j\}, 1, Q^i_g) - \tilde{W}_g(\{i,j\}, 0, Q^i_g)\right\}.$$

Marital gains:

$$\varphi_g(\{i,j\}) = \tilde{U}_g\left(\{i,j\}, Q^i_g\right),$$

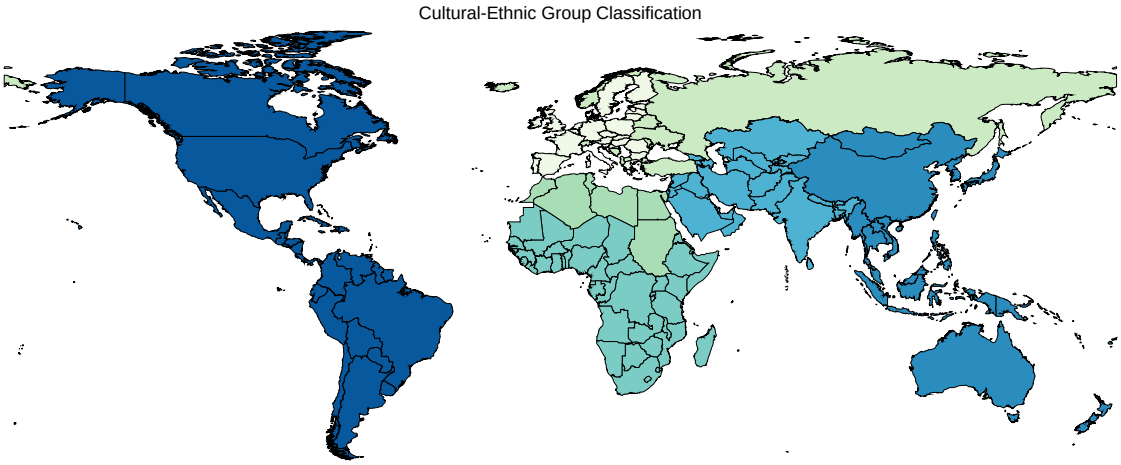
where

$$\varphi_g(\{i,j\}) = \ln \frac{\mu_g\left(\{i,j\}, Q^i_g\right)^2}{\mu_g\left(\{i,0\}, Q^i_g\right)\mu_g\left(\{0,j\}, Q^i_g\right)}.$$

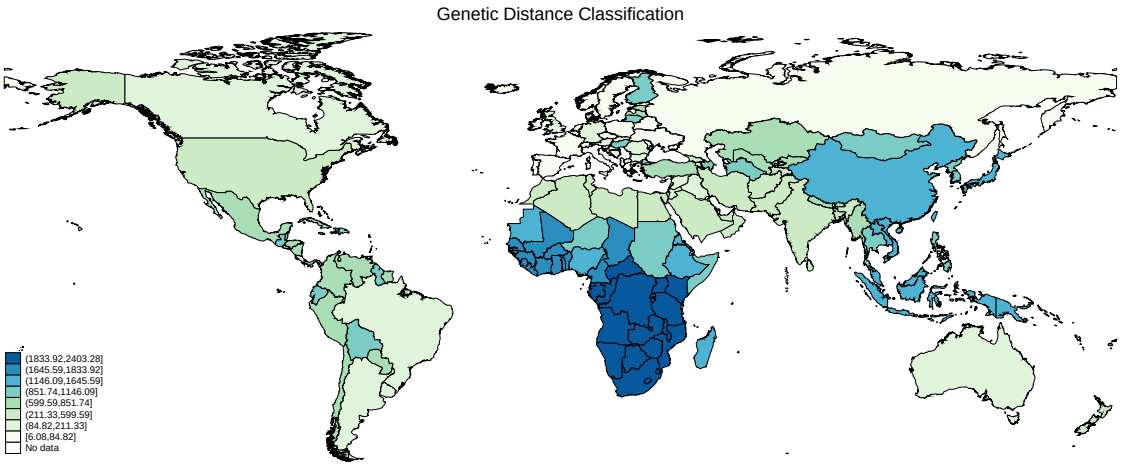
B.4 Additional Figures and Tables

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

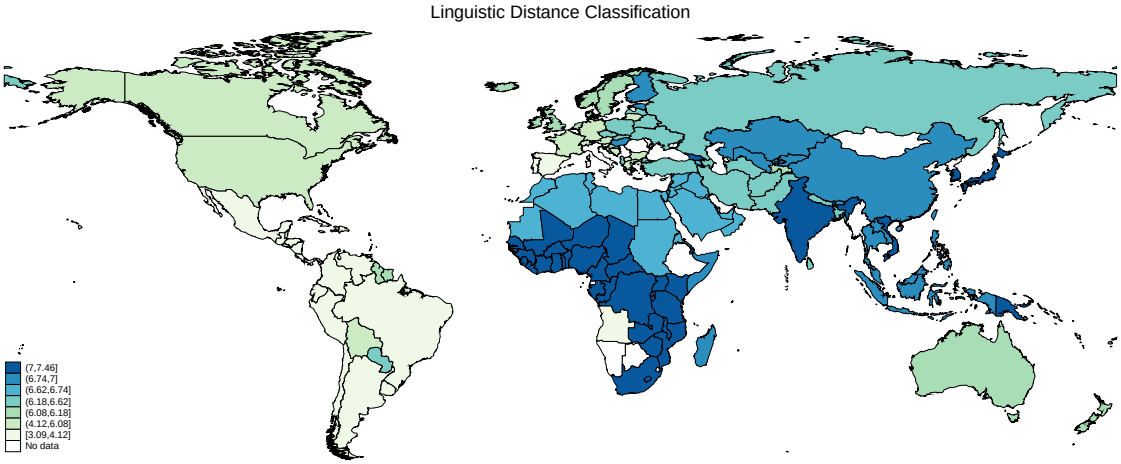
(a) Cultural Ethnic Group Classification



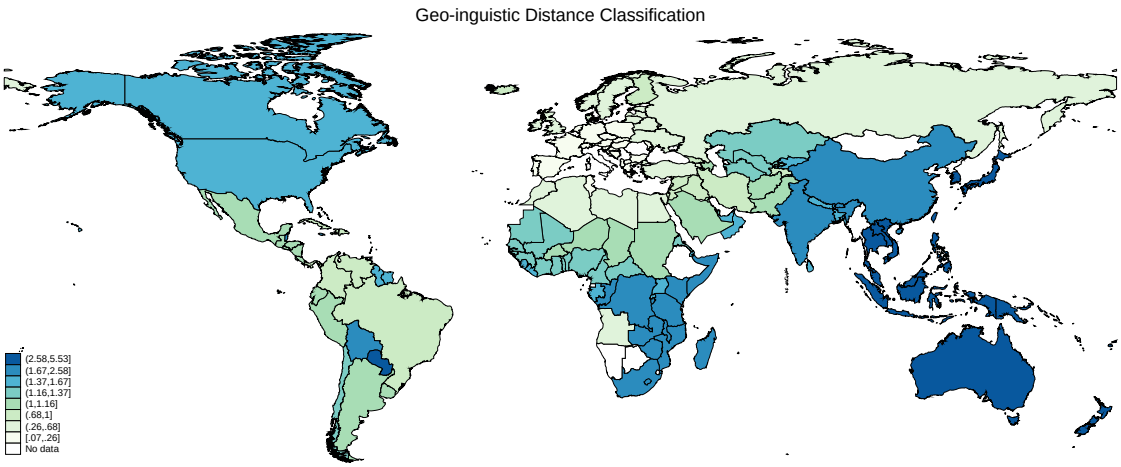
(b) Genetic Distance Classification



(c) Linguistic Distance Classification



(d) Geo-Linguistic Distance Classification



Note: Representation of our ethnic-group classification and countries' cultural distances with respect to Italy. The genetic distance measures the degree of genealogical relatedness between two populations and it is associated to the time elapsed since two populations' last common ancestors. The metrics for genetic distance is based on the genetic tree classification, as described in Cavalli-Sforza and Piazza (1994). It is defined based on the *coancestry coefficients*: the heterozygosity index, i.e. the probability that two alleles from a given locus, selected at random from two populations, will be different. This implies that the higher the genetic distance between two populations, the longer the separation period between them and the larger the difference in vertical cultural characteristics. Data on genetic distance metrics are available thanks to Spolaore and Wacziarg (2009). Measures for ethnolinguistic distance are provided by Melitz and Toubal (2014b) and Egger and Toubal (2016), and are based on the language tree classification proposed by Lewis et al. (2009).

Table B.2: Cultural-Ethnic Group Classification of Migrants' Countries of Origin

Cultural-Ethnic Group	(Share %)	Countries
Europe 27 - E	26.63	Austria, Belgium, Bulgaria, Cyprus, Croatia, Czech Republic, Denmark, Estonia, France, Finland, Germany, Greece, Hungary, Ireland, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, United Kingdom, Slovakia, Slovenia, Spain, Sweden
Other Europe - O	23.61	Albania, Andorra, Belarus, Bosnia and Herzegovina, Estonia, Iceland, Isle of Man, Liechtenstein, Kosovo, Macedonia (FYROM), Republic of Moldova, Monaco, Norway, Russian Federation, San Marino, Vatican City State, Serbia and Montenegro, Switzerland, Turkey, Ukraine, Vatican City State
North Africa - N	14.40	Algeria, Egypt, Libyan Arab Jamahiriya, Marocco, Tunisia
Other Africa - 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
South-West Asia - W	8.66	Afghanistan, Saudi Arabia, Armenia, Azerbaijan, Bahrain, Bangladesh, Bhutan, United Arab Emirates, Georgia, India, Islamic Republic Of Iran, Iraq, Israel, Kazakhstan, Kyrgyzstan, Kuwait, Lebanon, Maldives, Nepal, Oman, Pakistan, Qatar, Syrian Arab Republic, Sri Lanka, Tajikistan, Palestinian Territory, Turkmenistan, Uzbekistan, Yemen
East Asia - S	8.61	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
America - M	8.69	Antigua and Barbuda, Argentina, Bahamas, Barbados, Belize, Plurinational State of Bolivia, Brazil, Canada, 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, United States, Uruguay, Venezuela

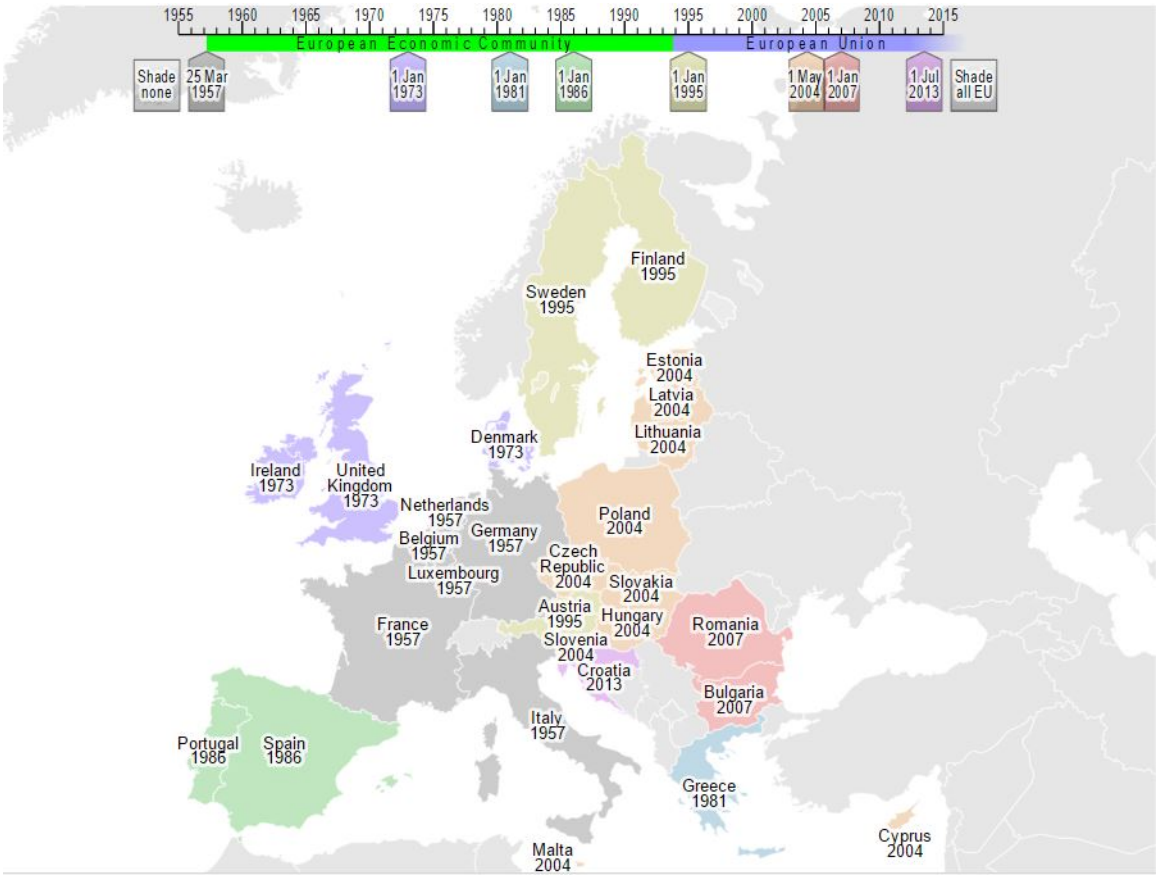
Note: *Source:* ISTAT Classification of foreign countries by cultural-ethnic groups. Ethnic group shares are computed as percentages over the aggregate migration inflows from 1995-2010.

Appendix C

The Price of Citizenship: The Effect of EU Enlargement on Marital Matching in Italy

C.1 Additional Figures and Tables

Figure C.1: Complete Timeline of the EU Enlargements from 1957



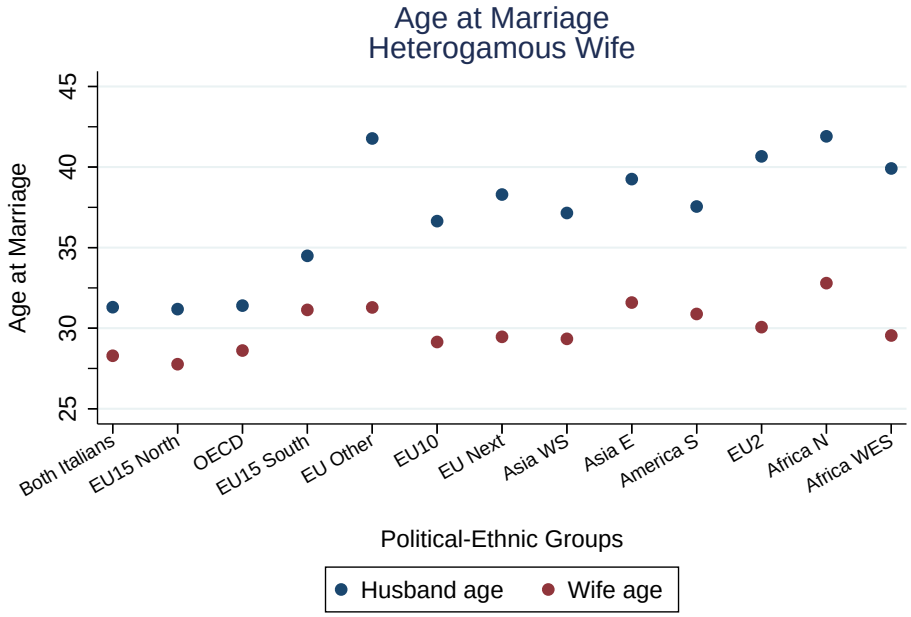
Source: https://ec.europa.eu/neighbourhood-enlargement. The figure reports the evolution over time of the process of EU enlargements from 1957.

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

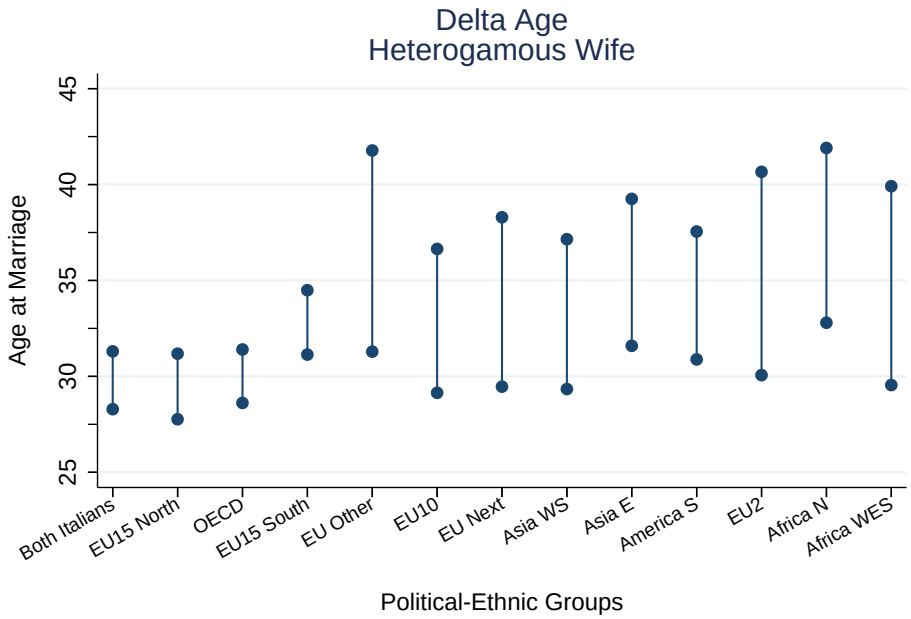
Ethnic Group	Countries
EU15 German	Austria, Belgium, Denmark, Finland, Germany, Ireland, Luxembourg, Netherlands, United Kingdom, Sweden
EU15 Latin	France, Greece, Portugal, Spain
EU10	Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovakia ,Slovenia
EU2	Bulgaria, Romania
EU Next	Albania, Bosnia and Herzegovina, Croatia, Iceland, Kosovo, Macedonia (FYROM), Serbia, Montenegro, Turkey
EU Other	Andorra, Belarus, Isle of Man, Liechtenstein, Norway, Monaco, Republic of Moldova, Russian Federation, San Marino, Ukraine, Vatican City State
North Africa	Algeria, Egypt, Libyan Arab Jamahiriya, Marocco, Tunisia
Other Africa	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
South-West Asia	Afghanistan, Saudi Arabia, Armenia, Azerbaijan, Bahrain, Bangladesh, Bhutan, United Arab Emirates, Georgia, India, Islamic Republic Of Iran, Iraq, Israel, Kazakhstan, Kyrgyzstan, Kuwait, Lebanon, Maldives, Nepal, Oman, Pakistan, Qatar, Syrian Arab Republic, Sri Lanka, Tajikistan, Palestinian Territory, Turkmenistan, Uzbekistan, Yemen
East Asia	Brunei Darussalam, Cambodia, China, Democratic People's Replica of Korea, Philippines, Jordan, Indonesia, Lao Pepople's Democratic Republic, Malaysia, Mongolia, Myanmar, Singapore, Taiwan, Thailand, East Timor, Vietnam, Fiji, Kiribati, Marshall Islands, Federated States of Micronesia, Nauru, Palau, Papua New Guinea, Solomon Islands, Samoa, Tonga, Tuvalu, Vanuatu
South America	Antigua and Barbuda, Argentina, Bahamas, Barbados, Belize, Plurinational State of Bolivia, Brazil, Canada, 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
Other OECD	Australia, Canada, Japan, Republic of Korea, New Zealand, Switzerland, United States

Note: *Source:* ISTAT Classification of foreign countries by political-ethnic groups.

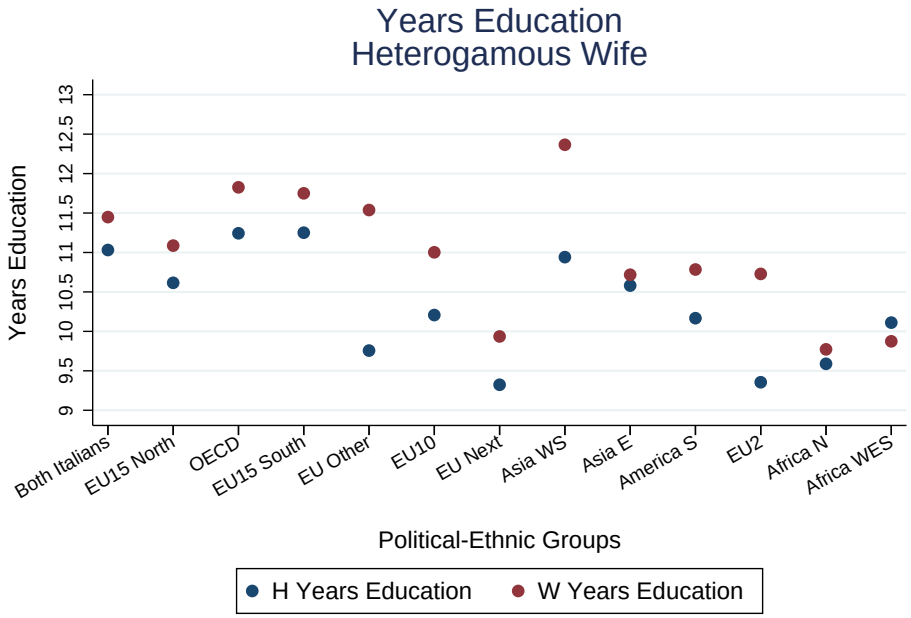
C.1.1 Age and Education



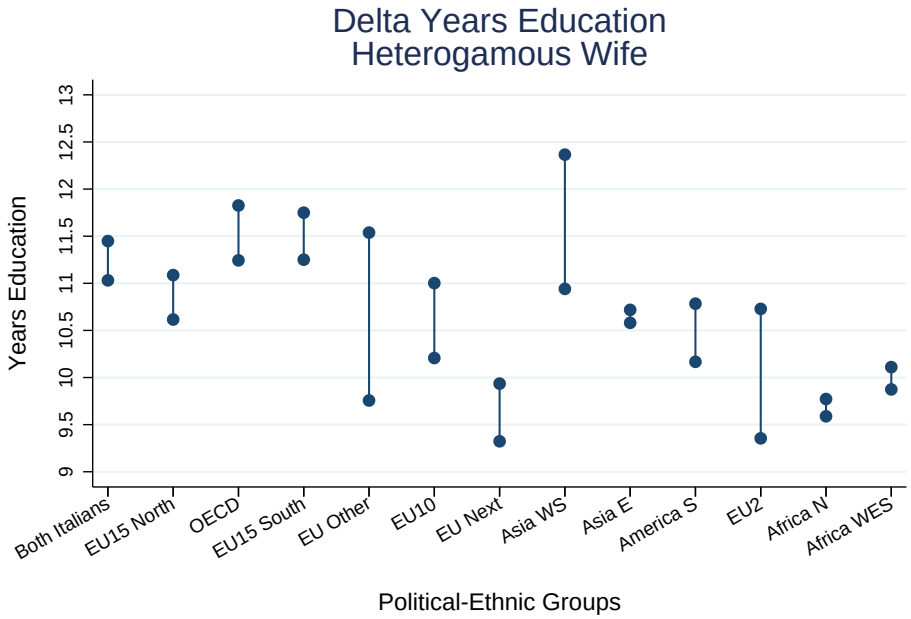
Note: Rank wrt average GDPpc 2010 by Ethnic-Group



Note: Rank wrt average GDPpc 2010 by Ethnic-Group



Note: Rank wrt average GDPpc 2010 by Ethnic-Group



Note: Rank wrt average GDPpc 2010 by Ethnic-Group