There's More to Marriage than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations*

Jérôme Adda† Bocconi University

Paolo Pinotti‡ Bocconi University

Giulia Tura§ EUI

February 14, 2019

Abstract

We relate intermarriage formation to cultural distance between spouses and legal status acquisition motives in the context of a marital matching framework. We empirically estimate the effect of legal status acquisition on gains from marriage exploiting registry data on the universe of marriages in Italy and exogenous variation in immigrants' legal status, as determined by the EU enlargement process. Other things equal, access to legal status reduces by half the probability of immigrants intermarrying with natives. Building on this evidence, we develop and structurally estimate a multidimensional equilibrium model of marriage and separation allowing for trade-offs between cultural distance, legal status, and other socio-economic spouses' characteristics, where individuals match on observed and unobserved characteristics. The model allows to separate and quantify the effect of legal status from cultural affinity between different areas of origin.

Keywords: Intermarriages, marital matching, separations, legal status, cultural distance
JEL Classification: J11, J12, J15.

*The data on marriages and separations used in this paper have been access through the Laboratory for the Analysis of Elementary Data (ADELE) at ISTAT, in compliance to the laws on the protection of statistical confidentiality and of personal data. We are the solely responsible for the results and the opinions expressed in this paper, which do not constitute official statistics. We thank Francesco Fasani for sharing the data on the 2002 amnesty. We are grateful to Alfred Galichon, Aloysius Siow and participants in seminars at many universities and conferences for helpful comments and discussion.

†Dept. of Economics, Bocconi University. email: jerome.adda@unibocconi.it.

‡Dept. of Social and Political Sciences, Bocconi University, Dondena, and fRDB. email: paolo.pinotti@unibocconi.it.

§Dept. of Economics, European University Institute. email: giulia.tura@eui.eu.

1 Introduction

Immigration is deeply transforming the structure of families in destination countries. In the wake of increasing migration flows, intermarriages between natives and foreigners account for a large and increasing share of total marriages. In 2014, the share of intermarriages over total marriages reached 14 percent in France, 12 percent in the United States and Germany, and 9 percent in Italy (OECD, 2017). This phenomenon has important implications for the cultural assimilation of foreign spouses and, also, of the second generations – because of parental cultural transmission.

Equilibrium models of marital matching à-la-Becker (Becker, 1973, 1974) predict positive assortative mating along cultural lines, provided that similarity in the ethnic origin of spouses enhances the household production function. Within this framework, the higher the cultural distance between the spouses, ceteris paribus, the lower the incidence of intermarriages relative to homogamous marriages. Indeed, intermarriages are often interpreted as a sign of integration of minorities (see, e.g., Gordon, 1964; Meng and Gregory, 2005; Constant and Zimmermann, 2008; Algan et al., 2012).

However, the intermarriage rate depends on another critical factor that works in the opposite direction: legal status acquisition motives. In most destination countries, foreign spouses of native citizens enjoy immediate access to residence and work as well as a privileged path to citizenship. These benefits increase the marital surplus in intermarriages relative to that in homogamous marriages, so foreigners may be willing to trade off cultural distance for legal status. More generally, we interpret migrants' marital decisions as the result of a tension between positive mating along cultural lines, to enhance complementaries within the household production function, and preferences to intermarry with natives, driven by the desire to acquire legal status.

From an empirical perspective, it is generally difficult to disentangle the effect of legal status acquisition motives and cultural distance, as they would jointly affect the relative prevalence of intermarriages and homogamous marriages. In addition, legal status may be correlated with several personality traits that are relevant for marital matching decisions, so marriage rates according to the legal status of the foreign spouse do not tell us about the effect of legal status acquisition motives. To overcome this identification challenge, we use the successive enlargements of the European Union (EU). Between 2004 and 2007, 12 countries, accounting in total for about 80 million inhabitants, joined

the EU.[1] As a consequence, their citizens acquired permanent legal status, including the right to free movement and work, within all member countries – apart from transitory restrictions to employment that were imposed in some EU countries. Therefore, the EU enlargement eliminated legal status acquisition motives to intermarry for new EU citizens – but not for other immigrants. This episode constitutes an ideal research design for separately identifying the effect of legal status acquisition and of cultural affinity from other socio-economic characteristics.

Intuitively, by looking at marriages involving new EU citizens, the decline in marriages with natives and the increase in separations reveal how much those citizens valued legal status. The incentives for natives are somewhat different as they would still grant legal access to any foreign spouse they marry. If new EU citizens are less keen to marry them, they have to turn to other types of spouses, perhaps more culturally distant or with different characteristics. This trade-off then reveals the value of cultural affinity and of differences in terms of, e.g., age and education. As the enlargement was staggered, we observe marriage choices and separations in three different environments – before, throughout, and after the enlargement period – with varying payoffs. We exploit this intuition by analysing registry data on the universe of marriages and separations in Italy over the period 1998-2012. This allows us to look at individual decisions involving 3.6 million of marriages and over 200 thousand separations before and after the enlargements. We approach this problem in two separate and complementary ways.

First, we show using difference-in-differences and event study techniques that the enlargements of the EU had a profound effect on mixed marriages and separations as soon as information became known. The probability of intermarrying with native males decreases by half, ceteris paribus, for females from new EU countries relative to females from other countries after (the announcement of) the EU enlargement.[2] During the same period, mixed couples composed of natives and citizens from the enlargement coun-

[1] These countries were Eastern European countries (Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, and Slovenia) plus two Mediterranean countries (Malta and Cyprus).

[2] Results are similar for intermarriages between native females and immigrant males, though the latter account for only a minor share of all intermarriages. They are robust to including or excluding the (estimated) number of undocumented immigrants; to including non-registered partnerships, as measured by cohabiting couples in the Census; and to controlling flexibly for the number of potential matches by nationality.

tries experienced an increased separation rate. We also document substitution patterns across national groups in the choice of spouses. The sharp decline in marriages between natives and Polish citizens following the entry of Poland into the EU favored mixed marriages involving Romanian spouses. However, this marriage market unraveled after a few years when Romania entered the EU. This in turned favored mixed marriages involving spouses from other Eastern European countries.

A second contribution of this paper is to develop and structurally estimate a multidimensional equilibrium model of both marriages and separations, allowing for trade-offs between cultural distance, legal status, and other socio-economic spouses' characteristics, such as age, educational attainment and wealth. In contrast to the first part of the analysis, we jointly model marital sorting and dissolution choices into one coherent framework and we achieve parameters' identification by exploiting both observed matching patterns and separation rates over time. Another important aspect is that the model allows us to explicitly deal with the equilibrium nature of the marriage market, to consider counterfactuals and to estimate the importance of general equilibrium effects (the "spillover effects" emphasized by Choo and Siow, 2006a,b) in response to the EU enlargements.

We depart from the standard workhorse model in this literature derived in Choo and Siow (2006b) in several ways. First, our model is multidimensional, a feature shared with the recent literature (Dupuy and Galichon, 2014; Ciscato et al., 2015). Second, we extend the model to a multi-market framework and we exploit variation over time (before and after the EU enlargement) and across geographical areas in population vectors to achieve model identification.[3] Third, we propose a model where individuals match on both observables and unobservables, and we depart from the literature by considering an environment where marital surplus is specific to a given couple in order to realistically describe marriage and separation choices over time at the couple level. The existing models in the literature rest on individual additively separable random shocks over observable characteristics of the potential spouses, which rules out any complementarity between unobserved characteristics (Chiappori and Salanie, 2016).[4] We introduce a match specific

[3] In a similar fashion, Chiappori et al. (2017) investigate changes in the returns to education on the US marriage market in the aftermath of World War II by exploiting variation across cohorts.

[4] A direct consequence of the separability assumption is that if two men belong to the same group, and their respective partners belong to the same group, the two potential matching allocations in the market

random component which allows for sorting on unobserved characteristics on both sides of the market. Moreover, our specification also allows for complementarity between observables and unobservable characteristics, by relaxing the separable linearity regarding the idiosyncratic shock.[5] The cost of this departure is to lose closed-form tractability, but we show that a simulation based method is feasible to estimate this richer model, where we use a triple difference method for identification for both marriages and separations of different cohorts married before or after the EU enlargements, across different geographical areas and different set of characteristics including the country of origin.

Our analysis provides evidence on the value of legal status, the absence of which leads to a reduction in the marriage surplus ranging from about 80 percent in the North of Italy to 70 percent in the South of Italy – where informal work makes legal status less valuable. This effect dwarfs that of education on marriage surplus, and is similar to that of home ownership.

These results contribute to our knowledge of the consequences of legal status for immigrants' behavior in receiving countries. Previous work has examined the effect of legal status on labor market opportunities (Amuedo-Dorantes et al., 2007; Lozano and Sorensen, 2011) and criminal activity (Mastrobuoni and Pinotti, 2015; Pinotti, 2017; Fasani, 2018), and consumption (Dustmann et al., 2017). However, the implications of legal status for marriage formation have been largely neglected with the notable exception of Azzolini and Guetto (2017), who document the negative effect of the EU enlargement on the number of intermarriages using synthetic control methods. Relative to this previous work, we provide a theoretical and empirical (structural) framework that allows us to study not only the effect of the EU enlargement on the number of marriages, but also on their characteristics and on cross-nationality substitution in marriage markets. In addition, we also estimate the impact of legal status acquisition on separations.

The estimation also produces valuable insights into the role of cultural affinity across different nationality groups. This measure of affinity is obtained from revealed preferences through marriage choices. We show stark differences in the marriage surplus,

lead to the same total surplus. This allows to translate the individual matching problem into a matching along observables.

[5] In a different approach, Galichon and Salanié (2015) maintain the additive separability but relax the multinomial logit structure in Choo and Siow (2006b) by allowing for correlations between shocks. Chiappori et al. (2017) allow for heteroskedasticity in the error term.

which decreases with cultural distance. We externally validate the marriage penalty associated with cultural affinity using standard measures of cultural distance recognized in the literature, such as genetic or language distance. We show that our measure, although correlated with these measures offers a new perspective on cultural affinity. Religious or linguistic distance appear to play a minor role, while genetic distance accounts for a bit short of half of the variation in our cultural affinity measures derived from the marriage surplus. Perhaps surprisingly, the cultural dimension in marriage has been overlooked, despite its relevance for various social and economic behaviours and the strong persistence of cultural traits across generations (Bisin and Verdier, 2000; Bisin et al., 2004; Fernández et al., 2004; Doepke and Zilibotti, 2008; Alesina et al., 2013).[6] We use our model to evaluate a counterfactual policy, which would grant legal status to all migrants in the country. Such policies have been enacted in the past in countries such as Italy, France, Spain or in the US. These policies foster integration through access to the legal labor market. We show that such measures reduce mixed marriages in favor of homogamous ones. Paradoxically, such measures may prevent integration of migrants along cultural lines, and may have intergenerational consequences.

The paper is organized as follows. Section 2 describes the institutional background and the characteristics of the EU enlargements to East European countries. Section 3 describes the data and empirical strategy and presents the empirical results concerning the effect of legal status on gains to marriage and on separations. Section 4 develops and structurally estimates the multidimensional equilibrium model of the marriage market and subsequent separations. Finally, Section 5 concludes.

2 Institutional background

2.1 Italian migration policy

Immigration to Italy is a relatively recent phenomenon. The number of official foreign residents increased tenfold between 1990 and 2017 – from 500 thousand to 5 million.

[6]The previous literature on marriage has focused on the sorting mechanisms along age (Choo and Siow, 2006b; Choo, 2015); income and body mass index (Chiappori et al., 2012); educational attainment (Chiappori et al., 2009, 2017; Gayle and Shephard, 2019); and personality traits (Dupuy and Galichon, 2014).

Figure 1 shows the composition of immigrant population by area of origin. Just less than a third of all foreign residents – 1.5 million – come from another country within the EU. By virtue of the Schengen agreement, all EU citizens may freely circulate and work in Italy.

The access of immigrants from all other countries is instead regulated by a rigid quota system. The main pathway to an official stay is work permits, issued each year by the national government. The so-called *Decreto Flussi* ("Flows Decree") sets stringent limits on the number of permits available by type of contract and province, and applications must be backed by job offers from prospective employers in Italy. Applicants who eventually obtain a work permit are allowed to reside in Italy for the duration of the job; their spouses and children are entitled to a residence permit for "family re-unification". If the job contract is terminated, however, the foreign worker has 6 months in which to find a new job, otherwise (s)he must leave the country. The application for Italian citizenship requires 10 years of continuous (legal) residence. Overall, Italian migration policy is quite restrictive. For instance, Italy ranks third out of 12 EU countries in the Strictness of Immigration Policy index produced by the Fondazione Rodolfo Debenedetti.[7]

However, such restrictions do *not* apply to foreign spouses of Italian citizens. They enjoy immediate access to residence and work in Italy, and can apply for citizenship after 2 years of marriage (as opposed to 10 years of continuous residence in Italy).[8] Intermarriage with natives thus constitutes an attractive gateway to residency, work, and citizenship in Italy. The same rights to residence and work in Italy – but not the preferential path to citizenship – also apply to foreigners from non-EU countries married to non-Italian EU citizens.

Figure 2 shows the share of foreigners over total residents and the intermarriage rate, separately for foreign females and males, over the period 1996-2013. The intermarriage rate is defined as the share of marriages contracted between a foreign and a native spouse over the total number of marriages contracted in a given year. This graph conveys two main facts. First, native males and foreign females tend to intermarry more than native females and foreign males; our theoretical model in Section 4 will account for this fact.

[7]The index is described at http://www.frdb.org/page/data/categoria/international-data/scheda/inventory-of-migration-policies-1990-2005/doc_pk/11028.

[8]The period required to apply for citizenship increases to 3 years if the spouses reside outside Italy after marriage, whereas it is shortened by half if the couple has children.

Second, intermarriage rates increase in parallel with the share of foreign residents until the second half of the 2000s, but diverge thereafter. Indeed, the growth in the share of foreigners over total residents accelerates starting in 2007, whereas intermarriage rates flatten out, especially for immigrant females, during the same period. Both these changes in trends coincide with the admission of millions of Eastern Europeans to free circulation in all EU countries, including Italy.

2.2 *The EU enlargement*

The EU is an economic and political partnership of 28 countries. This configuration is the result of subsequent enlargements that are still ongoing, as several countries are negotiating admission conditions. It was instituted by the Maastricht Treaty on November 1, 1993, and consisted of 12 countries.[9] It was not until a decade later that Eastern European countries were also admitted to the EU. The first round of the enlargement became effective on May 1, 2004, and involved ten countries (*EU*2004 henceforth): Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovakia and Slovenia. Three years later, on January 1, 2007, Romania and Bulgaria also joined the EU (*EU*2007 henceforth).

Both rounds of the enlargement were preceded by complex and lengthy negotiations. Figure 3 shows the timing of final decision stages. The EU Council – the executive branch of the Union – agreed upon the 2004 enlargement on December 2002; the timing for the 2007 enlargement is similar, the EU Council taking the final decision on December 2005. Therefore, uncertainty regarding the outcome of both enlargements unraveled at the beginning of the previous calendar year. Our empirical analysis will allow for anticipation effects during such period.[10]

In principle, citizens from new member countries acquired the right to reside and work in all other countries of the EU; in practice, many countries in the latter group adopted transitory regimes imposing barriers against new EU citizens from Eastern Europe. For instance, Germany eliminated entry quotas for Polish workers only in 2011.

[9] The first 12 countries were Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, and the United Kingdom; Austria, Finland, and Sweden joined the EU two years later.

[10] A detailed description of the overall process can be found here: https://ec.europa.eu/neighbourhood-enlargement.

Similarly, until 2014 Bulgarians and Romanians needed to acquire a special permit to work in several EU member states – Austria, Belgium, France, Germany, and the United Kingdom.

By contrast, Italy largely adhered to the principle of free circulation, maintaining only mild restrictions to employment of new EU citizens in some sectors of the economy. In 2004 and 2005 the Italian government set a cap of 20 and 79.5 thousand, respectively, to the number of employees from new EU countries. On the other hand, there were neither restrictions to residency nor to non-dependent working activity (i.e., self-employment or entrepreneurship). As for the 2007 enlargement, the government also lifted restrictions to dependent employment in most economic sectors.

3 The Effect of the EU enlargement on marriages and separations

Figure 4 shows the evolution of marriages and separations between natives and immigrant spouses, by area of origin, during the period 1998-2012 (all series are indexed to 1 in 1998). Figure 4a shows that intermarriages between natives and new EU citizens were on an upward trend until the announcement of EU council decisions (i.e., in 2002 and 2005, respectively) and declined sharply immediately after. In contrast, marriage rates between natives exhibit a steady and slow decline over that period. Figure 4b shows the dynamics of separation, the formal act that starts the divorce proceedings (as explained in Section 3.3, at least three years of legal separation are required for obtaining divorce). Separations between new EU immigrants and natives exhibit abnormal spikes after the EU enlargements, to decline immediately after.

This preliminary evidence is consistent with legal status acquisition reducing incentives to marrying with natives for immigrants from new EU countries. At the same time, migration inflows from such countries – notably, after the EU enlargement – confound the identification of the effect of legal status in Figure 4. We next illustrate our strategies for identifying changes in the propensity to intermarry and in the hazard rate of separation, controlling for the total number of immigrants by area of origin.

3.1 *Marriages: data and methodology*

The marriage rates depicted in Figure 4a do not take into account changes in the number of available matching opportunities in the market – notably, the large increase in the number of foreigners from new EU countries. An increase in mixed marriage rates could occur either because those marriage become more desirable or, mechanically, because there are more available individuals of a given nationality. To take into account the change in the pool of available men and women, we follow the methodology introduced by Choo and Siow (2006b). We rescale the number of total marriages between males from country x and females from country y by the geometric average of single males and females to calculate the gains to marriage :

$$\Phi(x,y) = \ln \frac{marriages(x,y)^2}{singles(x) \cdot singles(y)}. \tag{1}$$

Intuitively, a higher number of observed marriages between type-x males and type-y females, relative to the number of potential matches between the same types, signals higher gains to marriage.[11] The gains to marriage $\Phi(x,y)$ in equation (1) will be the main dependent variable in our empirical analysis. We focus on heterosexual marriages in which at least one of the spouses is Italian, and compute $\Phi(x,y)$ across cells defined by the country of origin of the foreign spouse, year of marriage, and Italian province of residence of spouses.[12]

We compute $\Phi(x,y)$ in equation (1) across country of origin-province-year cells exploiting rich registry data on the universe of marriages formed in Italy and on the number of foreign residents. These data were made available by ISTAT through its Laboratory for Elementary Data Analysis (ADELE).[13]

[11] Choo and Siow (2006b) show that $\Phi(x,y)$ equals the systematic component of gains from marriage in the unique equilibrium of a random utility, unidimensional matching model with transferable utility between spouses. Section 4 extends the model to a multidimensional setting in which spouses match on multiple individual traits.

[12] Italian provinces correspond to level 3 of the EU Nomenclature of Territorial Units for Statistics (NUTS). In Census year 2001 there were 99 provinces, with average and median population of 551 and 377 thousand inhabitants, respectively.

[13] Access to the data for research purposes is regulated by ISTAT through an open application procedure. Authorized researchers use the data from work stations located in secure rooms and isolated from the Internet as well as from any input or output devices. The output of analysis is made available only at the end of the project upon inspection by ADELE officers. For more information, see http://www.istat.it/en/information/researchers/laboratory-for-elementary-data-a.

Vital statistics registries provide detailed information on the universe of marriages celebrated each year in Italy between 1998 and 2012 – more than 3.6 million marriages (see Panel A of Table 1). Using information on spouses' country of origin, reported in the marriage records, we compute the numerator of the marital matching surplus $\Phi(x, y)$ in equation (1).[14] Alternatively, following Choo and Siow (2006a) we will measure the numerator of $\Phi(x, y)$ by the sum of legal marriages and cohabiting couples not in a legal relationship, available from the decennial population Census.

Measuring the number of singles (i.e., the denominator of the marital matching surplus $\Phi(x, y)$ in equation 1) raises two main challenges. First, the exact number of residents by country of origin and marital status is available only from the decennial Census. We thus count the singles in 1991, 2001, and 2011 by the number of unmarried individuals aged between 18 and 60; Panels C and D of Table 1 show the number of observations and summary statistics for years 2001 and 2011, respectively. We then interpolate these data outside Census years using data on official migration inflows. The exact procedure is described in Appendix A.

Second, neither the Census nor other official migration statistics include data on unofficial immigrants. However, neglecting unofficial migrants from the denominator of (1) would bias the estimated gains from marriage upward for all immigrants except new EU citizens in the post-enlargement period. This would lead us, in turn, to over-estimate the reduction in the gains from marriage for such groups after the EU enlargement, as we discuss in detail in Section 3.2 and in Appendix A. To avoid this problem, we account for the presence of undocumented immigrants drawing on (i) administrative data on applications for amnesty by irregular immigrants in year 2002, and (ii) yearly estimates of the total number of irregular immigrants present in Italy during the period 1991-2013.

The general amnesty of 2002 is the largest one ever enacted in Italy, granting a temporary residence and working permit to about 700 thousand irregular immigrants – a third of the total immigrant population at that time.[15] We obtained the universe of application

[14] In addition to country of origin, marriage records report spouses' date of birth, educational attainment, previous marital status, employment status and professional level, municipality of residence at the time of marriage, and province of residence after marriage. In Section 4 we use this information to investigate trade-offs in spouses' characteristics.

[15] All immigrants that had been residing for at least three months in Italy and were (unofficially) employed at the time of the amnesty were eligible for a 2-year residence permit. The working condition was certified by the employer. Importantly, regularized immigrant workers did not obtain permanent legal

files reporting, among other things, the gender, age, marital status, country of origin, and Italian province of residence of each applicant. Based on this information, we compute the number of singles (i.e., unmarried individuals aged between 18 and 60) among unofficial immigrants in year 2002, by country of origin and Italian province of residence. We then impute the total (estimated) number of unofficial immigrant singles in all other years to country of origin-province cells based on the distribution of applicants from amnesty across the same cells. Appendix A provides the details of this procedure and the data used to implement it. The total number of singles in each cell equals the sum of singles among official and unofficial migrants in such cell. We investigate the robustness of results to excluding (imputed) unofficial immigrants from the computation of singles.

We estimate the following equation for the gains to marriage $\Phi_{IT,y,t,p}$ between native males ($x = IT$) and females from country of origin y, in province p, and year t:

$$\Phi_{IT,y,t,p} = \beta(newEU_y \times postEU_{y,t}) + D_y + D_p + D_t + \varepsilon_{IT,y,p,t}, \tag{2}$$

where $newEU_y$ is a dummy for females from $EU2004$ or $EU2007$ countries; $postEU_{y,t}$ is a dummy variable for the years after each enlargement, i.e. $postEU_{y,t} = 1$ from 2004 (2007) onwards for females from $EU2004$ ($EU2007$) countries; D_y, D_p, and D_t are dummy variables for each country of origin, province, and year, respectively; finally, $\varepsilon_{IT,y,p,t}$ is an error term summarizing the effect of other determinants of gains to marriage. We estimate an analogous equation for the gains to marriage between native females and males from country of origin x, $\Phi_{x,IT,p,t}$.

The coefficient β in equation (2) captures the relative change in gains from intermarrying with natives after the EU enlargement between immigrants from new EU countries and other countries, respectively. Given that immigrants may acquire legal status (or even Italian citizenship) for other reasons than admission to the EU, β provides a lower bound, in terms of magnitude, to the effect of legal status acquisition. Put differently, it represents the "intention-to-treat" effect of legal status.

status – as it was the case, instead, for new EU citizens. Devillanova et al. (2017) provide additional details.

3.2 *The effect of the EU enlargement on marriage formation*

3.2.1 Main results

Figure 5 plots the gains to marriage with natives, as measured by $\Phi(x, y)$ in equation (1), for different groups of immigrants during the period 1998-2012.[16] The announcements of both the 2004 and 2007 enlargements bring a decline in gains to marriage between new EU citizens and natives, with the exception of marriages between *EU*2004 males and native females. However, the latter amount to less than 100 intermarriages per year, so gains from such marriages may be imprecisely estimated.

We quantify the changes in Figures 5 by estimating the difference-in-differences specification in equation (2); results are presented in Table 2. Column (1) reports the baseline estimated coefficient for marriages in which the husband is native, and column (5) reports the estimated coefficient for marriages in which the wife is native. Admission to the EU brings a 0.7 log point decrease in gains to marriage between native males and females from new EU countries, and an even larger decrease – twice as much – in gains to marriage between native females and males from new EU countries. These estimates are unaffected when interacting province fixed effects with, respectively, year and country of origin fixed effects (columns 2 and 6). The same is true when weighting observations by the province population (columns 3-4 and 7-8). The results in columns (4) and (8) of Table 2 imply that, keeping constant the number of singles in the market, province$\times$year and province$\times$origin fixed effects, admission to the EU decreased the number of marriages with native males and females by 51 and 75 percent, respectively.

Figure 6 plots the estimated coefficients and confidence intervals obtained when interacting a full set of year fixed effects with separate dummies for marriages between natives and, respectively, immigrants from *EU*2004 and *EU*2007 countries (see also the Appendix Table B3). The first two graphs confirm that marriages between native males and females from new EU countries drop markedly after the latter obtain legal status. Interestingly, the 2004 enlargement also induced a strong increase in marriages between native males and *EU*2007 females. This is consistent with native males substituting *EU*2007

[16] Figure B3 in the Appendix shows the evolution of the different components of $\Phi(x, y)$, namely the number of intermarriages and the number of immigrant singles. Table B1 reports the complete list of countries in each group.

for $EU2004$ females when the announcement of the 2004 enlargements lowers (expected) gains to marriage for the latter group.

Marriages between native females and males from $EU2007$ countries also decrease strongly after the announcement of the 2007 enlargement, while the evidence is less clear for males from $EU2004$ countries. As we already mentioned, however, there are very few marriages between native females and males from $EU2004$ countries. Finally, Figure 6 documents the absence of differential trends before the announcement of either enlargement.

3.2.2 Robustness

In Table 3 we investigate the sensitivity of our baseline estimate to relaxing two restrictions imposed by the model of Choo and Siow (2006b). First, their empirical framework implies that, for constant gains to marriage, percent changes Δ_x and Δ_y in the number of male and female singles, respectively, would induce a percent change $\sqrt{\Delta_x \Delta_y}$ in the number of marriages. In column (1), we relax this restriction by controlling on the right-hand side of equation (2) for the (log) number of single females at the denominator of equation (1); notice that the number of (native) single males is absorbed by the set of province$\times$year fixed effects. Analogously, in column (5) we control for the log number of single males.

Second, the ratio $\Phi(x, y)$ in equation (1) is invariant to the number of singles of different origins who could also be potential matches.[17] In practice, however, the propensity of foreigners to intermarry with natives likely reflect the availability of other singles, particularly from their same country of origin. Indeed, the period after the EU enlargement witnessed large inflows from new EU countries, so the estimated decrease in gains from marrying with natives could reflect greater availability of singles from the same country of origin. To separately identify this effect from the effect of legal status acquisition, in column (2) of Table 3 we control for the log number of single males from the same country of origin as the female spouse, in each province and year. Analogously, in column (6) we control for the log number of females from the same country of origin of the male spouse, in each province and year.

[17] This issue is similar to the Independence of Irrelevant Alternatives in multinomial logits and stems from the functional form assumptions in Choo and Siow (2006b).

Finally, in columns (3) and (7) we include all control variables at the same time, and in columns (4) and (8) we further interact them with country of origin fixed effects. In all these specifications, the effect of interest remains strongly statistically significant and qualitatively similar to the baseline estimate in Table 2. In particular, in the most stringent specification (columns 4 and 8) the effect is very similar – about -0.85 log points – between marriages formed by native males and females, respectively.

In Table 4 we provide additional robustness checks. In columns (1) and (5) we re-estimate equation (2) after excluding unofficial immigrants – whose number is imputed – from the denominator of the gains from marriage in (1). Consistently with the discussion in Section 3 and Appendix A, the estimated (negative) effect of the EU enlargement on gains to marriage between new EU citizens and natives becomes larger when including only official immigrants in the computation of the number of singles. However, this larger effect comes from underestimating the number of singles for all immigrant groups except new EU citizens in the post-enlargement period. For this reason, we prefer to include unofficial immigrants throughout the analysis.

We next investigate whether legal status acquisition produced "real" effects on the number of native-immigrant couples actually living together or, rather, just a substitution from official marriages into cohabitations of unmarried couples. To this purpose, in columns (3) and (7) of Table 4 we replicate the analysis after adding cohabiting, unmarried couples to married couples at the numerator of the gains from marriage. The sum of married and unmarried couples living together is available from the 2001 and 2011 Census.[18] If the effect of interest was entirely (or partly) driven by substitution of official marriages with cohabitations, we should observe a zero (or weaker) effect when measuring gains from marriage using the total number of couples living together. Instead, the estimated effect of interest is still negative and significant, and its magnitude is only slightly affected, suggesting that our results are not explained by substitution from official marriages into cohabitations. Therefore, the enlargement produced important effects on the actual number of couples formed between natives and new EU citizens.

In even columns we re-estimate all regressions in odd columns controlling for the

[18]We do not have, unfortunately, information regarding the starting date of the cohabitation. In order to retrieve a flow dimension, we thus select cohabitations with migrants permanently residing in Italy in the last five years before the two Census, i.e. from 1997 to 2001 for the 2001 Census, and from 2007 to 2011 for the 2011 Census.

log number of unmatched singles interacted with country of origin fixed effects (i.e., the most stringent specification in columns 4 and 8 of Table 3). Consistently with the evidence in this previous table, the results are qualitatively unaffected and the point estimates becomes much more similar between marriages formed by native males and females, respectively.

Results are also robust to re-estimating the effect assuming that the whole of Italy is a single marriage market; to leaving out one country of origin at a time; to accounting for sample selection potentially induced by the fact that there are no marriages in some cells (so the logarithm in (1) is not defined); to alternative one-way and two-way clustering of standard errors. These results are available upon request. Finally, in Table B4 of the Appendix we provide indirect evidence that online marital matching was not a relevant phenomenon in Italy during our period of interest. In particular, we show (i) that the (log) number of marriages observed in Italy mirrors the distribution of singles across province-nationality-year cells, and (ii) that the coefficient of interest does not vary with the diffusion of broadband Internet by region and year. We next move to the evidence on separations.

3.3 Separations: data and methodology

ADELE provides access to rich administrative data on separation records from civil courts' registries. Such data cover the universe of separations in Italy from 1998 to 2012, providing detailed information (among other things) on separation proceedings, date and place of marriage, and date and place of birth for both spouses. Separation is the formal act that starts the divorce proceeding. Until 2015, a minimum period of 3 years of legal separation was required for eventually obtaining divorce. For this reason, separations provide a more accurate representation of the timing of marital instability.

We matched marriage and separation records on the exact date of marriage and exact date and place of birth of both spouses, which are available in both registries. The combination of these characteristics allows an exact one-to-one matching for 99.5% of couples, and we discard the remaining fraction. The final dataset includes over 200 thousand separations over the period 1998-2012 (see Panel B. of Table 1).

We follow each married couple over time until potential separation. Following the popular approach by Cox (1972), we assume that the hazard rate is a log-linear function

of observable covariates and an arbitrary baseline hazard common to all couples:

$$h(d, Z) = \lambda(d) \exp(\gamma' Z), \tag{3}$$

where $h(d, Z)$ is the hazard rate of separation after d years of marriage, conditional on the vector of covariates Z; $\lambda(d)$ is the baseline hazard; and γ is the vector of covariates' coefficients. Letting the baseline hazard unrestricted, we can estimate γ by partial maximum likelihood. This approach remains very tractable while allowing for considerable flexibility in the baseline hazard rate $\lambda(d)$. In particular, we allow the baseline hazard to vary by province, year, and between homogamous native and mixed couples, respectively. In this way, the baseline hazard absorbs, among other things, the systematic change (if any) in the risk of divorce for mixed couple after two years of marriage – i.e., when the foreign spouse can apply for Italian citizenship. As for our main effect of interest, the effect of the EU enlargement is identified off differential changes in the hazard rate of separation between new EU citizens and other immigrants, before and after the enlargement.

3.4 The effect of the EU enlargement on separations

Table 5 shows the estimated effect of the EU enlargement on the hazard rate of separations, modeled as in equation (3). The specification in columns (1) and (5) includes on the right-hand side the dummy *newEU*, its interaction with dummies for the period after the enlargement, and calendar year fixed effects. We include separate interactions for the year after the enlargement and for the following years, respectively, as the decrease in gains to marriage should produce an immediate increase in separations immediately after the enlargement. This is indeed the case. The hazard rate of separation between native males and females from new EU countries increases by 23 percent in the year immediately after the enlargement. The effect is lower (15 percent) and not statistically significant for marriages between native females and males from new EU countries.

Marriages formed after the enlargement are particularly solid. This is shown in columns (2) and (6), in which we interact *newEU* with dummies for marriages formed before and after the enlargement period – respectively, 1998-2002 and 2007-2012. Other things equal, marriages formed between natives and immigrants from new EU countries after the enlargement period have a 35 percent lower probability of breaking up relative to other marriages. All results are confirmed when including all interactions with *newEU* in the

same specification and allowing for both calendar year fixed effects and different baseline hazards by year of marriage (columns 3 and 7). Finally, they are also robust to controlling for the number of potential matches from the same country of origin as the foreign spouse (columns 4 and 8)

4 Multidimensional marital matching model

The previous section has shown how the successive EU enlargements have affected mixed marriages and their stability. The results point to an important role for legal status. Moreover, the patterns of substitution between spouses from different origins over time suggests that cultural affinity plays an important role in marital sorting.

In addition to cultural origins, individuals also match on other observable characteristics, such as age and education. Figure 7 shows the average age (top row) and education (bottom row) of spouses for different types of marriages during the period 1998-2002 (i.e., before the EU enlargement). The first graph shows the average age of spouses when they both come from the same country, by area of origin. Husbands are on average older than wifes, by two to four years. The gap increases considerably for intermarriages formed by native husbands and foreign wifes (second graph), while the opposite is true for intermarriages formed by native wifes and foreign husbands (third graph). Turning to education, wifes are on average more educated than husbands (fourth graph in Figure 7). The gap widens for intermarriages between native husbands and foreign wifes, while it is reduced for intermarriages between native wifes and foreign husbands.

This shows that sorting in the marriage market is multidimensional, and individuals match by trading off different characteristics of their potential spouses. To the extent that individuals prefer on average to match with younger and more educated spouses, this evidence is consistent with immigrants trading off spouse's age and education for legal status acquisition.[19] The model introduced by Choo and Siow (2006b) that we used in Section 3 offers a transparent identification of the effect of the enlargement. It is however limited in describing the multidimensional nature of the marriage market. To address

[19]The Appendix Table B2 provides additional evidence in this respect by estimating a difference-in-differences specification for spouses' age and education in intermarriages with spouses from new EU and other countries, respectively, before and after the EU enlargement.

this issue, we develop a multidimensional equilibrium model of both marriage and separations, allowing for trade-offs between cultural distance, legal status, and other socio-economic spouses' characteristics, such as age, educational attainment and wealth. Our model allows us to describe jointly marital sorting and dissolution choices into one coherent framework, by considering an environment where marital surplus is specific to a given couple and changes over time. Thanks to the introduction of a match specific random component, our model allows for sorting on observable and unobservable characteristics of spouses on both sides of the market. In this respect, we considerably depart from the existing models in the literature, which rest on individual additively separable random shocks over observable characteristics of the potential spouses.

4.1 Setup

We consider a transferable utility (TU) marital matching model in a frictionless – but local – marriage market. The TU assumption captures the idea that spouses implicitly transfer utility between each other in absence of transaction 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 matches in the market. Moreover, in absence of frictions in the market, we assume that any individual has complete and costless information about all subjects present in that market and their observable characteristics.

We extend the model to a multi-market framework along two dimensions of interest. First, We exploit the longitudinal structure of the data focusing on two time periods. Let t denote the time dimension where $t \in \{t_1, t_2\}$, with t_1 indicating the period before the EU enlargements (1998-2002) and t_2 the period after (2007-2012). Second, we assume that each province p corresponds to a specific local marriage market.[20] Given the temporal and geographical heterogeneity, we end up with $TP = 20$ separate and independent marriage markets.

Each marriage market is a two-sided market, with a population of men denoted by $i \in I$ and women $j \in J$, residing in a given province p at time t. Men and women are

[20] We focus on $P = 10$ different provinces. Those provinces are Milano, Torino, Bologna, Genova, Firenze, Perugia, Rome, Napoli, Bari and Palermo. These provinces were chosen to get the largest ones and in order to have a range of provinces across northern, central and southern parts of the country.

heterogeneous along a vector of relevant attributes for the marriage market. In particular, each man $i \in I$ is characterized by a vector of observable attributes $x_i \in X$, while woman j is characterized by a vector of attributes $y_j \in Y$. In our empirical application x_i and y_j represent a bundle of personal and socio-economic characteristics,

$$x_i = \{x_i^A, x_i^{Ed}, x_i^{Orig}, x_i^H\}, \quad y_i = \{y_i^A, y_i^{Ed}, y_i^{Orig}, y_i^H\},$$

comprising respectively age, educational attainment, geographic area of origin, and wealth (home-ownership). Home-ownership is not recorded on marriage certificates, but this information appears in the Census, together with the other characteristics in x_i or y_i. This allows us to compute the conditional probability of home-ownership that we denote $\pi_t^W(\bar{x}_i)$ and $\pi_t^W(\bar{y}_j)$, where $\bar{x}_i = \{x_i^A, x_i^{Ed}, x_i^{Orig}\}$ and $\bar{y}_i = \{y_i^A, y_i^{Ed}, y_i^{Orig}\}$ are the set of characteristics excluding home ownership. In our estimation method, we use those probabilities to integrate out housing when computing marriage rates.

Given I men and J women, $I \times J$ matches are potentially observed in each market p at each point in time t. A *matching* defines who is matched with whom and who remains unmatched. Specifically, a matching is a measure $\mu_{tp}(i,j)$ over the $I \times J$ space, such that $\mu_{tp}(i,j) = 1$ if man i is matched with woman j from the reference population in market p at time t, and zero otherwise. By considering a one-to-one matching framework, any individual in the market can be matched to only one person, the matching measure thus satisfies the following feasibility constraints:

$$\sum_J \mu_{tp}(i,j) \leq 1 \quad \forall i \in I, \qquad \sum_I \mu_{tp}(i,j) \leq 1 \quad \forall\, j \in J. \tag{4}$$

In addition, we define $\mu_{tp}(i,0) = 1$ if man i is single and similarly $\mu_{tp}(0,j) = 1$ if woman j is single in the same reference population.

4.1.1 Marriage surplus and optimal matching problem

Let $\tilde{\Phi}_{tp}(i,j)$ denote the joint utility generated by assigning man i to woman j, residing in area p in period t.[21] Denote $\omega_{tp}(x_i)$ and $\omega_{tp}(y_j)$ the outside option of remaining

[21] Because of the transferable utility nature of our matching framework, our interest lies in the estimation of the joint marital surplus rather than the spouses' individual utilities net of transfers, as they are not identified without additional intra-household observations.

single for man i of characteristics x_i and woman y_j, respectively. The joint marital *surplus* generated from the (i,j) marriage is thus, $\Phi_{tp}(i,j) = \tilde{\Phi}_{tp}(i,j) - \omega_{tp}(x_i) - \omega_{tp}(y_j)$ and a marriage takes place only if the joint net surplus is positive: $\Phi_{tp}(i,j) > 0$. Thus, under the assignment $\mu_{tp}(i,j)$, the total surplus generated within each market is equivalent to: $\sum_{IJ} \mu_{tp}(i,j)\Phi_{tp}(i,j)$. The *optimal* matching turns out to be the solution of the following welfare maximization problem over all potential matches in the market, including the outside options of singlehood:

$$\max_{\mu_{tp}(i,j)} \quad \sum_{IJ} \mu_{tp}(i,j)\Phi_{tp}(i,j) \tag{5}$$

subject to standard non-negativity and feasibility constraints in (4). The solution of the primal problem in (5) corresponds to a stable matching (Shapley and Shubik, 1971). 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 with other partners than remain in their current situation (i.e. absence of blocking pairs, Chiappori and Salanie, 2016). We solve the optimal matching (5) numerically, given the absence of closed form solution of our model. We detail this below.

4.1.2 Surplus specification

We characterize the joint utility generated within the marriage to include both systematic and idiosyncratic components. In particular, the joint marital surplus comprises four systematic components denoted by $\phi_p^k(x_i, y_j)$ with $k = \{A, Ed, Orig, W\}$, describing the effect of age, education, geographic area of origin, and wealth of a match between man i of characteristics x_i and woman j of characteristics y_j. In addition, the surplus depends on a time varying component $\beta_{tp}(x_i, y_j)$, which represents the value of legal status obtained through marriage, detailed below.

The surplus also depends on a match specific random component, $\varepsilon_t(i,j)$, which ensures that observationally identical agents, in equilibrium, will have different types of partners. By introducing an idiosyncratic component to the surplus generated by matching man i with woman j, we specifically allow for sorting on unobserved characteristics on both sides of the market. We interpret $\varepsilon_t(i,j)$ as a random love component between

man i and woman j, which is observed by the individuals, but unobservable from an econometric perspective.[22] We assume it follows an exponential normal distribution, with zero mean and variance σ_ε. In this respect, our model represents a departure from the literature originating with Choo and Siow (2006b), which relies on individual and separable idiosyncratic shocks over observed characteristics of mates only. Such an assumption leads to a tractable model but rules out complementarity between unobserved characteristics (Chiappori and Salanie, 2016). The surplus is represented as follows:

$$\Phi_{tp}(i,j) = \varepsilon_t(i,j) \cdot \beta_{tp}(x_i, y_j) \cdot \prod_{k=A,Ed,Orig,W} \phi_p^k(x_i, y_j) - \omega_{tp}(x_i) - \omega_{tp}(y_j). \tag{6}$$

Compared to the additive structure exploited in the existing literature, we allow the shock $\varepsilon_t(i,j)$ to enter multiplicatively in the surplus function. "Love" is therefore complementary to other observable characteristics.

We discretize each of the characteristics x_i^k and y_j^k and we allow for N_k groups, indexed by g and g' respectively. For age, we group individuals in six five-year groups, between the age of 22 and 50. We allow for two groups for education (more than high school or not) and two groups for wealth, whether the individual owns property or not. We assign individual cultural identities based on $N_{Orig} = 8$ geographic areas , namely Italy, EU15, EU2004/07, Rest of Europe, Asia, Africa, South-America and Other OECD countries. Denote $I_{x^k=g, y^k=g'}$ an indicator variable equal to one if the kth characteristic of both partners belongs to the group g and g' respectively. We parametrize the $\phi_p^k(x_i, y_j)$ components of the surplus as follows:

$$\phi_p^k(x_i, y_j) = \sum_{g,g'}^{N_k} \alpha_{p,g,g'}^k I_{x^k=g, y^k=g'} \quad , \quad k = \{A, Ed, Orig, W\} \tag{7}$$

where the parameters $\alpha_{p,g,g'}^k$ account for the strength of mutual attractiveness across spouses' observable characteristics. We allow $\alpha_{p,g,g'}^k$ coefficients to potentially vary across

[22] The idiosyncratic component might also be interpreted as an idiosyncratic non-monetary return from marriage. A different approach to rationalize the observed heterogeneity in marital sorting is to introduce some frictions in the market and model explicitly a meeting technology. The presence of frictions implies that any individual has imperfect and costly information about potential mates in the market. As stated in Chiappori and Salanie (2016), randomness on the meetings process guarantees that similar agents, in equilibrium, will have different types of partners.

provinces, but we restrict them to be constant over time, i.e. before and after the EU enlargement.

The marriage surplus also includes a potential penalty for not having legal access to the labor market, $\beta_{tp}(x_i, y_j)$. Natives and citizens of the EU countries have full access to the labor market, as well as their spouses. We normalize this coefficient to one for couples where both partners have legal access to the labor market when singles. This includes homogamous marriages between Italians or EU15 citizens before the enlargement, homogamous marriages of citizens from EU2004/07 after the enlargement, but also mixed marriages between Italians or EU15 citizens before the enlargement and with those from the new accession countries after the enlargement. For mixed couples where neither spouses have legal access to the labor market, we assume that the match carries a penalty $\beta_{tp}(x_i, y_j) = \underline{\beta}_p \leq 1$, which is potentially province dependent. It is possible that in the South of Italy, where informal work is more prevalent, legal access may not be worth much, so the penalty coefficient $\underline{\beta}_p$ may be closer to one, compared with provinces in the North. Finally, for couples where, as singles, one member has access to the legal market and the other not (e.g. a mixed marriage consisting of an Italian and an Asian spouse), we allow the surplus to vary by a factor $\beta_{tp}(x_i, y_j) = \bar{\beta}_p^{\male}$ or $\bar{\beta}_p^{\female}$ depending whether the Italian spouse is male or female.

Figure 8 gives a descriptive explanation of how we model the surplus related to legal status and cultural affinity. We abstract there from heterogeneity in age, education and wealth for expositional reasons. It also shows how identification is achieved. Data on marriages after the enlargement allow the identification of the cultural affinity parameters of the marriage surplus for marriages involving natives with foreigners belonging to the EU. For all other types of marriages, their proportion depends on both cultural affinity and the value attached to legal access. Hence, a cross-section does only allow the separate identification of a subset of parameters, related to cultural affinity. However, by also exploiting longitudinal variation, i.e. data before the enlargement, one can separately identify all of the parameters. For instance, the knowledge of the cultural affinity parameter between Italian men and women from the enlargement countries after the enlargement allows to identify the coefficient $\bar{\beta}_p^{\male}$ by looking at such marriages before the enlargement. Then in turn, this allows to disentangle the cultural affinity parameters for Italian men married to women from other (groups of) countries. Exploiting a typical difference-in-differences research design allows therefore to identify parameters relating

not only to marriages involving spouses from the enlargement countries but also for any other origin.

For identification purposes, we introduce the following coefficients' restrictions. For homogamous native couples, we impose that $\alpha^k_{p,1,1} = 1$ for all $k = \{A, Ed, Orig, W\}$. We also impose that the outside option of remaining single is the same across individuals, $\omega_{tp}(x_i) = \omega_{tp}(y_j) = \omega_{tp}$, but varies over time to account for differential trends in marriage rates. Hence the surplus for a couple of Italians, with low education, both aged 22 and with no housing, residing in province p at time t is equal to $1 - 2\omega_{tp}$. The outside option coefficient will be determined in the estimation so as to match the marriage rate for this group in province p at time t. The other coefficients $\alpha^k_{p,g,g'}$ are left free to vary to match the marriage rates for the other groups.

4.1.3 Divorce

We model divorce by introducing a dynamic structure at the level of the match specific random component, $\varepsilon_t(i,j)$. We rationalize it as an update of spouses' mutual love within marriage. We assume that $\varepsilon_t(i,j)$ evolves as a random walk:

$$\varepsilon_{t+1}(i,j) = \varepsilon_t(i,j) + \eta_{t+1}(x_i, y_j), \qquad \eta_{t+1}(x_i, y_j) \sim \mathcal{N}(0, \sigma^2_\eta(x_i, y_j)). \tag{8}$$

Negative surprises potentially trigger divorce choices, but we assume that only a fraction $1 - \delta$ of couples separate when the surplus become negative.

4.2 Estimation

Let $\boldsymbol{\theta}$ denote the vector of parameters describing the surplus of couples, defined above. We estimate the model parameters via a Method of Simulated Moments (MSM) estimator, matching a vector of observed moments from the data to a vector of theoretical moments predicted by the model, constructed through simulations. We denote by $\widehat{\Lambda}$ a vector of observed moments and by $\Lambda(\boldsymbol{\theta})$ the theoretical counterpart. The estimated surplus parameters are obtained as:

$$\hat{\boldsymbol{\theta}} = arg\min_{\boldsymbol{\theta}} \quad [\widehat{\Lambda} - \Lambda(\boldsymbol{\theta})]^\intercal \boldsymbol{W} [\widehat{\Lambda} - \Lambda(\boldsymbol{\theta})],$$

where $\boldsymbol{W}$ is a weighting matrix.

The vector $\widehat{\Lambda}$ contains observed marriage rates over observable characteristics $\bar{x}$ and $\bar{y}$ of spouses for any province p and period t, $\hat{\mu}_{tp}(\bar{x}, \bar{y})$[23], stacked with data moments characterizing divorce patterns of two different cohorts of married couples that we follow through time. Both cohorts were married before the enlargement, the first one is observed entirely before the enlargement, while the second is followed through it. Contrasting the separation patterns of the two cohorts allow us to evaluate the effect of the enlargement, separate from the propensity of separation of each type of couple.

Specifically, we construct separation rates for two cohorts, $c \in \{c_1, c_2\}$, that have been married for 2 to 4 years. The first cohort, c_1, consists of couples (of all origins) married in 1998-1999 and observed in 2000-01 when we see the status of their marriage. The second cohort, c_2, consists of couples married either in 2002-03 and followed to 2004-05 (couples involving either natives and/or migrants from the EU2004 enlargement) or married in 2005-06 and followed to 2007-08 (couples involving natives and/or migrants from the EU2007 enlargement). In the second cohort, citizens from the accession countries married to EU citizens may reconsider their initial marriage choice once they have legal access to the labor market granted through the enlargement of the EU. For each of these two cohorts, we regress an indicator variable D_{cij} equal to one if the (i, j) couple separate on a set of indicator variables coding for mixed marriages between Italians and new (or future) accession countries, between Italians and other countries of origin, for marriages involving a wife that is ten years younger or older than the husband and for marriages involving partners with the same level of education. Let Z_{cij} denote this set of regressors. The excluded category, captured by a constant, is for homogamous marriages. For each cohort c, we estimate the auxiliary parameters γ_c on the observed data from:

$$D_{cij} = Z_{cij}\gamma_c + u_{cij}. \tag{9}$$

[23] We recover matching patterns over observables $\bar{x}$ and $\bar{y}$ by to the following equality:

$$\mu_{tp}(\bar{x}, \bar{y}) = \sum_{IJ} \mu_{tp}(i, j) I_{tp}(\bar{x}_i = \bar{x}, \bar{y}_j = \bar{y})$$

where $I_{tp}(\bar{x}_i = \bar{x}, \bar{y}_j = \bar{y})$ is an indicator variable equal to one if man i of characteristics $\bar{x}_i = \bar{x}$ is matched to woman j of characteristic $\bar{y}_j = \bar{y}$, in province p at time t.

Therefore, the vector of observed moments is:

$$\widehat{\Lambda} = \{\hat{\mu}_{tp}(\bar{x}, \bar{y}), \hat{\gamma}_c\}, \quad t \in \{t_1, t_2\}, \;\; c \in \{c_1, c_2\}, \;\; p = 1, \ldots, P.$$

In total, once normalization restrictions are imposed, we have a total of 100 parameters to describe $(N_A.N_{Ed}.N_{Orig}.P)^2 = 92,160$ marriage types per period, as well as 14 moments describing separation choices.

Theoretical moments are computed as follow. We recover marital matching moments through simulations. For each period t and province p, we simulate N marriage markets of I men and J women. The distribution of individual characteristics in the simulated market parallels the distribution of observed traits for men and women, $m_{tp}(\bar{x})$ and $f_{tp}(\bar{y})$ at time t in province p that we observe in the data.[24] We use the observed probabilities of home ownership $\pi_t(\bar{x}_i)$ and $\pi_t(\bar{y}_j)$ to derive $m_{tp}(x)$ and $f_{tp}(y)$, the proportion of available men and women conditional on age, education, area of origin and wealth. We compute for all possible $I \times J$ matches the resulting surplus, and then solve for the optimal allocation in (5), using mathematical programming solver (Gurobi).[25]

In practice, some of the marriage types are rare, especially those involving minorities and for some educational groups. As in Dupuy and Galichon (2014), we assume that not all individuals interact together. Instead, we construct two pools where individuals meet. The first is exclusively composed of natives, whereas the second pool consists of natives and foreigners. We construct the second pool such that there are at least as many natives as foreigners, so that in principle all foreigners could find a native match. This procedure allows to have a boost sample of foreigners, and allows us to reduce the number of simulations. We assume that a fraction λ of Italians meet only Italians on the marriage market. The remaining fraction $1 - \lambda$ meet foreigners, and match according to both cultural, socio-economic characteristics and unobserved characteristics. We derive marital matching patterns, $\mu^{\theta}_{tp}(x, y)$, by recording as married any allocated match with a

[24] We recover population vectors, $m_{tp}(\bar{x})$ and $f_{tp}(\bar{y})$, in order to satisfy the following feasibility constraints:

$$\sum_Y \mu_{tp}(\bar{x}, \bar{y}) + \mu_{tp}(\bar{x}, 0) = m_{tp}(\bar{x}) \quad \forall x \in X, \qquad \sum_X \mu_{tp}(\bar{x}, \bar{y}) + \mu_{tp}(0, \bar{y}) = f_{tp}(\bar{y}) \quad \forall y \in Y. \tag{10}$$

A graphical representation of the distribution of observed traits by gender is reported in Figure B1.

[25] More information on Gurobi is available from http://www.gurobi.com/.

positive surplus, and consequently derive the theoretical distribution of single individuals, $\mu^{\boldsymbol{\theta}}_{tp}(x,0)$ and $\mu^{\boldsymbol{\theta}}_{tp}(0,y)$.

We consider $N = 8$ markets per province p, with size $I = J = 500$, half of them with foreigners. This amounts to consider 40,000 simulated males and females per period who potentially marry. Increasing N leads to more stable results as the sample size gets larger and more marriages between rarer types are observed. A larger N increases the computational time (in a linear way). The choice of I is undoubtedly arbitrary, but fixed both by considering computing time and stability of the estimates. Increasing I increases the computing time in a quadratic way. It also implies that each individual i samples more partners and therefore more shocks $\varepsilon_t(i,j)$, and hence have a larger chance of finding a suitable match. Holding the value of remaining single fixed (denoted ω_{tp} in equation 6), as I goes to infinity, everyone in the economy would get married. In other terms, the model is not scale invariant. However, once the value of remaining single is free to adjust with I, we observed that for a large enough I there were little changes in the overall allocation. Our choice of I reflects this fact, and allows for a tractable computing time.[26]

We construct the model counterpart of the separation rates by simulating two different cohorts. For the first cohort, the term $\beta_{tp}(x_i, y_j)$ remains constant through time, in such a way that, for this cohort only changes in the match specific random component account for separations. The second cohort sees a change of the surplus penalty term $\beta_{tp}(x_i, y_j)$ after their marriage decision in addition to the revision in their match specific random component. For a given choice of $\boldsymbol{\theta}$, we estimate the auxiliary parameters $\gamma^{\boldsymbol{\theta}}_c$ from (9) on simulated data, separately for each cohort c.

4.3 Estimated parameters

We first start with the fit of the model. We estimate 100 parameters in total. Out of those parameters, 35 of them pertain to the contribution of age to the marriage surplus and as many to the contribution of cultural affinity. The remaining parameters characterize the value of being single (11), the value of housing (6), the effect of legal status (6), education (3) and separations (4). Those parameters serve to describe 184,320 types of marriage rates across all characteristics, provinces and time period as well as 14 moments

[26] Such a setup can be rationalized with non-linear and convex search costs.

pertaining to separation rates. Figure 9 displays the fit regarding marriage patterns. The correlation between the observed and predicted marriage rates is equal to 0.83 before and after the enlargement. The prediction of the model regarding separation rates for cohorts followed before or during the enlargement process are displayed in Table 6. The model captures the general patterns of separation across cohorts and country of origin, although few parameters have a direct impact on separations in the model. Those are essentially four parameters describing the variance of the innovation of the idiosyncratic shock defined in equation (8) and two parameters capturing the fraction of couples who do not take into account the marriage surplus in their decision to stay married. The effect of the enlargement comes from the change in legal access, $\beta_{tp}(x_i, y_j)$, although that term serves also to match marriage choices before and after the enlargements.

We start by presenting the parameters related to age, which are displayed in Figure 10. The surplus is the highest for a couple where both partners are between 32 and 37, reflecting the higher marriage rates of this age category. The profiles are mostly hump-shaped, with a maximum when the spouses are the same age, or when the husband is slightly older. The lowest surplus component is for men aged 22 and women aged 50. The profiles are also asymmetric by gender. Couples where women are matched with a young man generate lower surplus, whereas the surplus profiles for men with younger spouses are flatter.

Table 7 displays the parameters pertaining to education and house ownership. We find mild assortative mating along education, with a higher surplus for couples with the same level of education. However, when education differs, the surplus is only 4 to 10 percent lower. We find much stronger effects for housing. When both partners are Italian, having one of them as homeowner increases the surplus by about 60 to 80 percent. The increase in the surplus is about 80 percent if both have property. We find effects of similar or slightly higher magnitude if one of the partner is non-Italian. The probability of owning property among non-natives is rather low, especially at younger ages. For non-natives from poorer countries (outside of the EU15 and OECD) the probability ranges between 10 to 30 percent, whereas for natives above the age of 30, the rate is at least two third. The results in Table 7 have two implications. First, given the asymmetry in the surplus associated with housing, mixed couples are more likely to include a native man and a foreign woman. Second, given the rate of ownership by groups, this aspect of the surplus tends to favor marriages between older native men (with home ownership)

and younger non-native women (without home ownership). The increase in the surplus compensates for part of the cultural and age differences among those couples.

Table 8 displays the effect of cultural affinity on the marriage surplus. The coefficient for homogamous native marriages (in the upper left corner) has been normalised to one. The lowest surplus is obtained for couples formed of an African or an Asian and a citizen from the original EU countries ($EU2004/07$). For those marriages, the surplus related to affinity is reduced by almost a half. In general, marriage surpluses are low in mixed couples when one of the partners is from either Africa or Asia. We find that the cultural affinity among homogamous marriages is much larger than within mixed ones. Mixed marriages are the outcome of assortativeness among other characteristics, including a strong couple love match, but are mainly achieved in our estimated model because of the value of legal access that we detail below.

The coefficients of the cultural affinity component of the surplus derive essentially from revealed preferences by observing the choice of spouses across periods and markets. It is interesting to relate this measure to other measures of cultural affinity commonly used in the literature: language, religion, values and genetics. Linguistic distance is based on the language tree classification, which groups languages into families based on perceived similarities: the lower the number of common nodes between two languages, the higher the distance between them. In a similar vein, religious distance originates from a tree-based representation of religions. Values measure distance in cultural norms, values and attitudes based on answers to the World Values Survey. Finally, genetic distance measures the probability that two alleles selected at random in two populations will be different: the higher the genetic distance between two populations, the longer they have been apart from each other, and the larger would be the difference in culture.[27] We show in Figure 11 how our measure of cultural affinity of the marital surplus relate to those alternative measures. Greater distance along cultural lines as measured in those four different ways is associated with a lower residual surplus from marriage. However, our measure of surplus is only weakly related to linguistic and religious differences. Those two measures only explain respectively 4 and 6 percent of the variation in the surplus. The World Value Survey measure explains 18 percent of the

[27] We refer to Spolaore and Wacziarg (2016) for an accurate description of these four measures of cultural distance.

variance, whereas genetic distance has the strongest correlation with our surplus measure and explains 44 percent of its variation. The political debate on migration often stresses the importance of religious or linguistic differences as factors preventing assimilation. In terms of marital choices, this does not appear to be the case.

We now turn to the effect of legal status on mariage surplus. The coefficients are displayed in the first panel of Table 9. The estimates of the penalty for lack of legal status ($\underline{\beta}$) show geographical heterogeneity. The absence of legal status leads to a reduction in the marriage surplus ranging from about 80 percent in the North of Italy to 70 percent in the South of Italy - where informal work makes legal status less valuable. Couples where one of the spouses grant legal status to the other have a much larger surplus, except in the case of a native wife and a foreign husband in the North, where there is a reduction in the surplus of about 20 percent. This parameter reflects the lower rate of mixed marriages where the wife is native, documented in Figure 2 and the subsequent analysis in Section 3.

Finally, we examine the effect of the enlargement on the formation of marriages between natives and citizens of the enlargement countries. From our model, the enlargement of the EU meant a large change in potential marital surplus for those couples. From Table 8, all else constant, a marriage between a native and a citizen from the New EU countries generates a slightly higher surplus than a homogamous native marriage, which is in term lower than the one for a homogamous marriage between two New EU citizens. Hence, after the enlargement, those types of mixed marriages are not many. However, before the enlargement, the total surplus takes also into account the value of legal access, which tends to largely favor those mixed marriages. This trade-off is pictured in Figure 12, which displays the surplus related to cultural affinity and legal status of women from the new EU countries, taking into account the origin of the spouse. After the enlargement, the surplus is highest for couples where both spouses are from the New EU countries, followed by couples involving an Italian husband and so on. The lowest surplus is for a couple with an Asian spouse. Before the enlargement, the surplus also takes into account the value of legal access to the labor market, and the ordering of the surplus is altered, with couples where an Italian husband or one from the original EU countries having the highest surplus.

4.4 Counterfactual analysis: generalized legal access

Many countries periodically grant amnesties to legalize irregular migrants. The leading example is the US Immigration Reform and Control Act in 1986, which granted legal status to about 3 million migrants. Related policies have been enacted in France and Spain. The goal of these amnesties is to integrate irregular migrants by granting them legal residence and allowing them to access labor markets legally. While many studies have focused on their effects on wages and employment (Cobb-Clark et al., 1995; Phillips and Massey, 1999; Amuedo-Dorantes et al., 2007; Lozano and Sorensen, 2011), or crime (Mastrobuoni and Pinotti, 2015; Baker, 2015; Pinotti, 2017), there is no evidence on their longer run impact on marriage markets. We address this question using our model. The variation we use to identify the model is closely related to such a policy, and therefore the counterfactual we construct relies much on the internal validity of the model.

We assess the effect of a generalized legalization on marriage markets using our estimated model and simulating a counterfactual where all foreigners present at the beginning of the period are granted legal status. We solve the model with these new features and compare to the baseline. In the counterfactual, we assume that $\underline{\beta} = \bar{\beta}^{\female} = \bar{\beta}^{\male} = 1$ for everyone in the economy. We report the results in Figure 13. The amnesty does little to the native marriage market, which increases by about 0.5 percent. There is a sharp decrease in mixed marriages, mirrored by a sharp increase in homogamous marriages among foreigners, subject of the amnesty.

While a legalization policy fosters the inclusion of foreigners into the labor market, it has also the unintended effect of reducing mixing in marriages. Given the strong intergenerational links in the transmission of cultural values within households, such policy thus prevents the assimilation of foreigners and their children and may have strong persistent effects.

5 Conclusion

In this paper, we exploit a natural experiment, the successive enlargements of the European Union, that shifted the incentives of some foreigners to marry natives. We show that the accession of the Eastern European countries changed profoundly the composition of mixed marriages. The migrants from the new European countries turned away

from natives, to marry more within their own communities. In turn, the natives changed their pattern of marriages towards migrants who did not (yet) have legal access to the labor market. Our analysis therefore stresses both the importance of legal access to the labor market in marriage decisions and how it can overcome differences in cultural affinity.

Our analysis relies both on reduced-form techniques and on a more structural approach. Both methods exploit not only marriage data across time and geographical areas, but also separation rates of couples followed over time. The identification is achieved through a difference-in-differences design in a multi-market framework. While the first part of the analysis has a transparent research design, it only allows to study one outcome at a time. The second part of the analysis puts more restriction on the data by parameterizing the surplus function, but allows to exploit jointly all data on marriages and separations to decompose the marriage surplus into a component that relates to the value of legal access or cultural affinity, among others.

We develop and structurally estimate a multidimensional equilibrium model of both marriage and separations, allowing for trade-offs between cultural distance, legal status, and other socio-economic spouses' characteristics, such as age, educational attainment and wealth. We extend the current literature on marriage models to a multi-market framework and to accommodate couple specific random components affecting marriage surplus. We show how such a model can be brought to data using simulation techniques.

We find that legal status is an important determinants of mixed marriages. The estimation of the model also allows to identify the cultural affinity aspect of marriage surplus. Cultural affinity is obtained through a revealed preference argument, by looking at substitution patterns in marriage choices. We find stark differences in preference towards spouses from various areas of origin. Interestingly, these preferences are correlated with genetic distance, but that measure explains only a fraction of the variability in preferences. Other measures such as religious or linguistic distance play very little role in explaining our measure of cultural affinity.

References

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

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

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.

Azzolini, D. and R. Guetto (2017). The Impact of Citizenship on Intermarriage: Quasi-Experimental Evidence From Two European Union Eastern Enlargements. *Demographic Research 36*, 1299–1336.

Baker, S. R. (2015). Effects of immigrant legalization on crime. *American Economic Review 105*(5), 210–213.

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.

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.

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 (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 (2017). Partner Choice, Investment in Children, and the Marital College Premium. *American Economic Review 107*(8), 2109–67.

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

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

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

Ciscato, E., A. Galichon, and M. Gousse (2015). Like Attract Like? A Structural Comparison of Homogamy Across Same-Sex and Different-Sex Households. *SSRN Working Paper*.

Cobb-Clark, D. A., C. Shiells, and B. Lowell (1995). Immigration reform: The effects of employer sanctions and legalization on wages. *Journal of Labor Economics 13*, 472–98.

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.

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

Devillanova, C., F. Fasani, and T. Frattini (2017). Employment of Undocumented Immigrants and the Prospect of Legal Status: Evidence from an Amnesty Program. *Industrial and Labor Relations Review (forthcoming)*.

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

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

Dustmann, C., F. Fasani, and B. Speciale (2017). Illegal Migration and Consumption Behavior of Immigrant Households. *Journal of the European Economic Association 15*(3), 654–691.

Fasani, F. (2009). Undocumented Migration. Counting the Uncountable. Data and Trends across Europe.

Fasani, F. (2018). Immigrant crime and legal status: Evidence from repeated amnesty programs. *Journal of Economic Geography 18*(4), 887–914.

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.

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

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

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

Guriev, S. M., B. Speciale, and M. Tuccio (2016). How Do Regulated and Unregulated Labor Markets Respond to Shocks? Evidence from Immigrants During the Great Recession.

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

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

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

Phillips, J. A. and D. S. Massey (1999). The new labor market: Immigrants and wages after irca. *Demography 36*(2), 233–246.

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

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

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

Figure 1: Number of foreign residents in Italy by area of origin (millions), 1994-2017

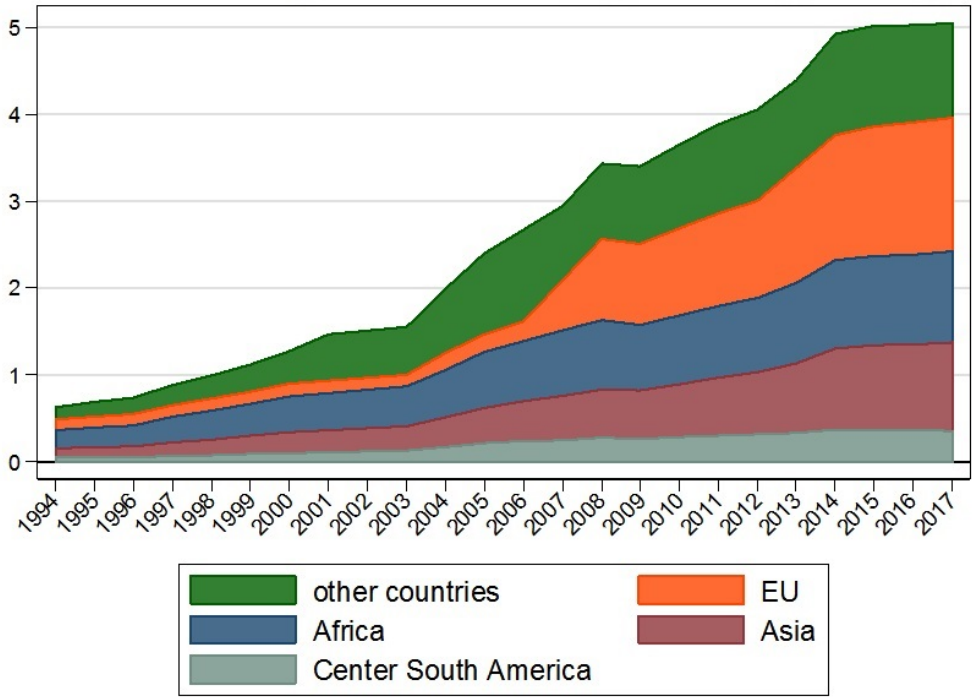
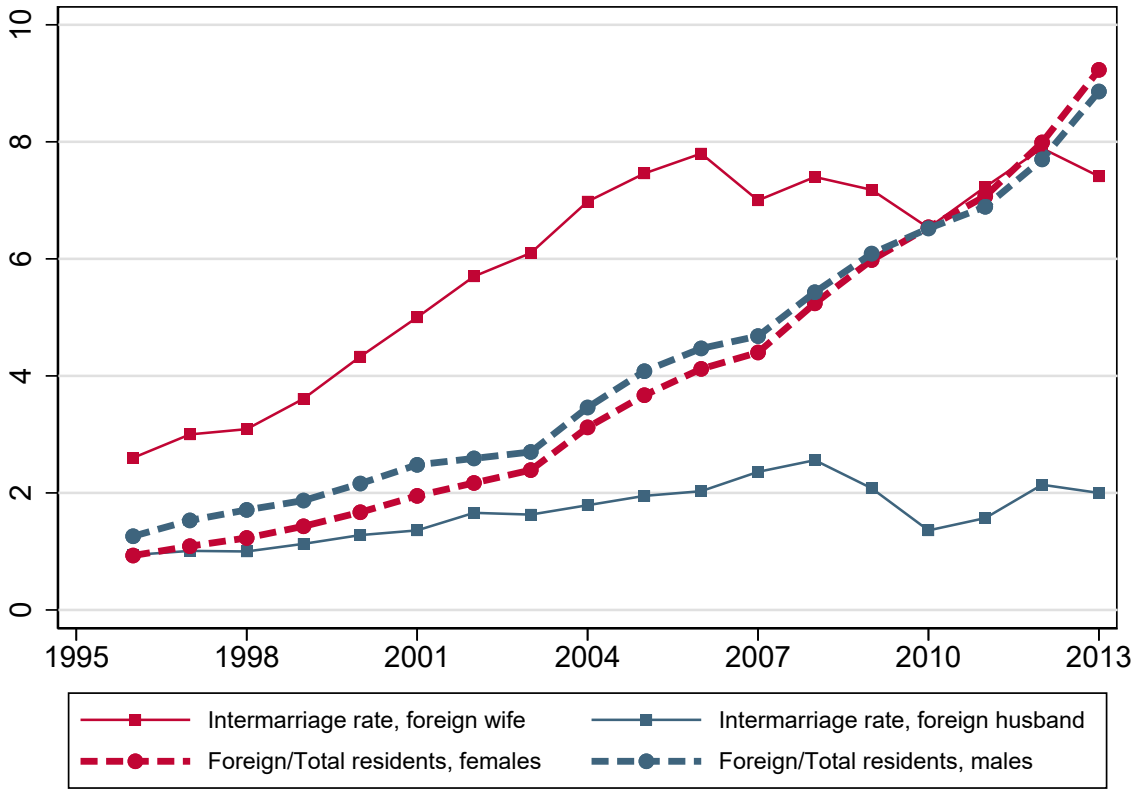


Figure 2: Intermarriage rate and ratio of foreign residents over total population, 1996-2013



This figure shows the evolution over time of the intermarriage rate in Italy, computed as the ratio between the number of intermarriages formed in a given year over the total number of marriages formed in the same year. The intermarriage rate is reported separately for intermarriages between an Italian husband and a foreign-born wife and an Italian wife and a foreign-born husband. The figure also shows the ratio of foreign over total residents, separately for females and males. *Source:* ISTAT, Marriage records from vital statistics registries and movements of foreign residents, 1996-2013.

Figure 3: Timeline of the 2004 and 2007 EU enlargements

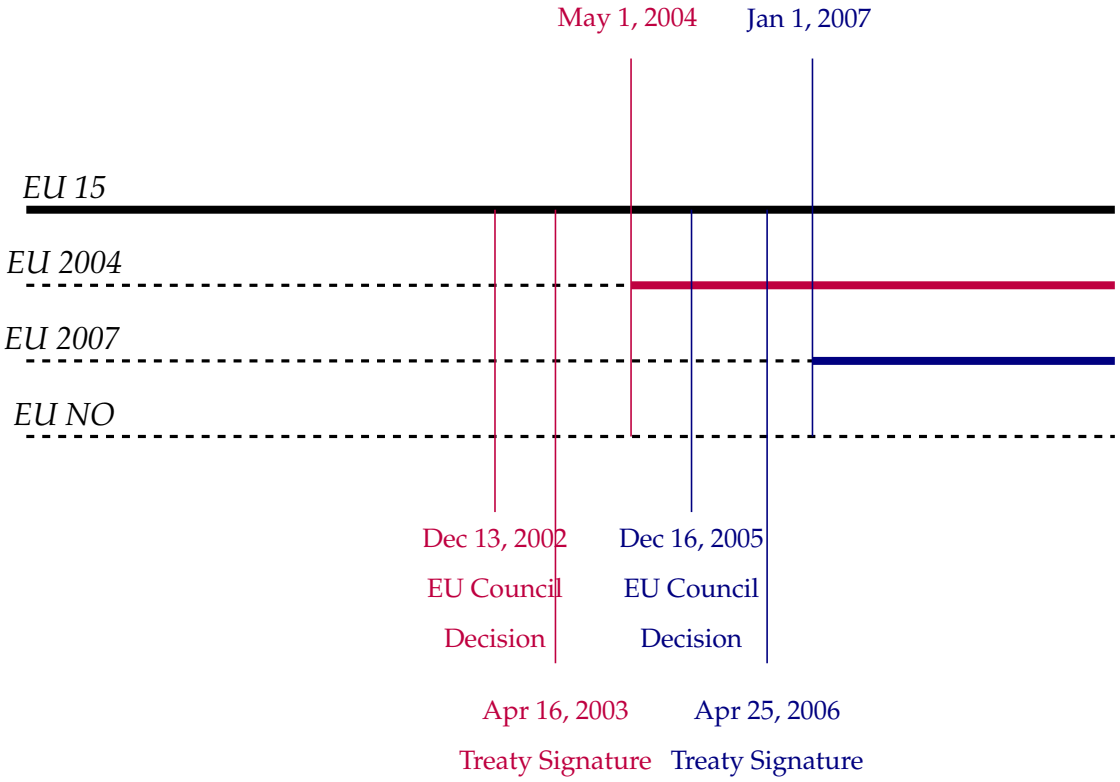
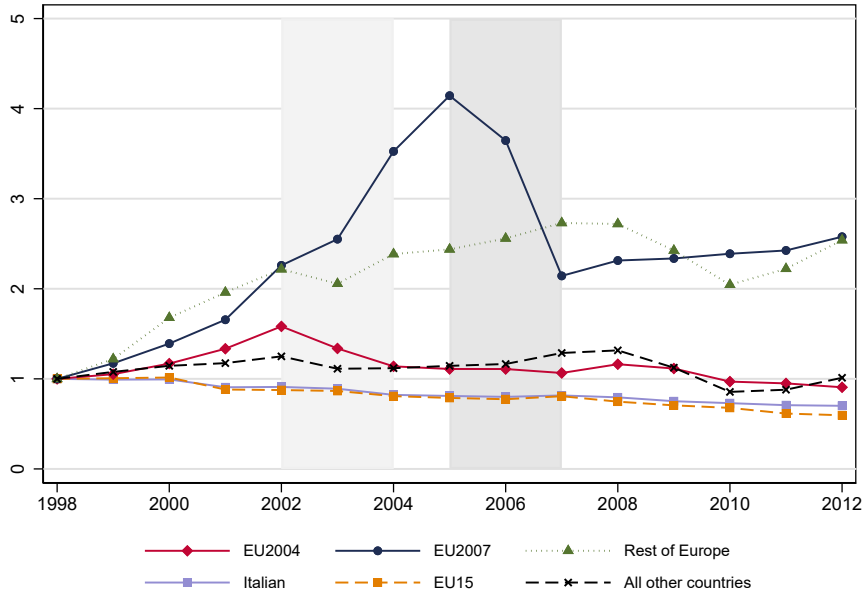
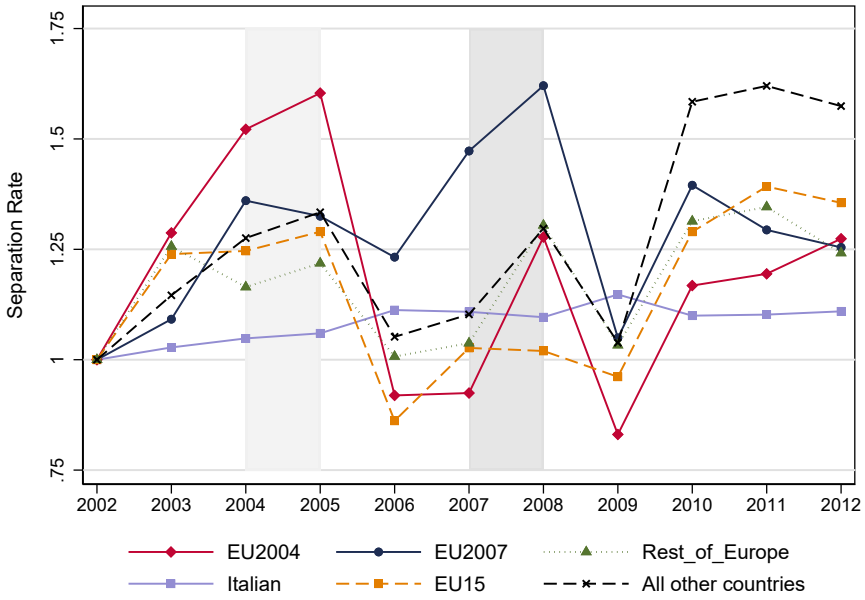


Figure 4: Marriages and separations, homogamous native couples and heterogamous couples formed by natives and immigrants, by area of origin of the foreign spouse

(a) Marriage rates (index=1 in 1998)



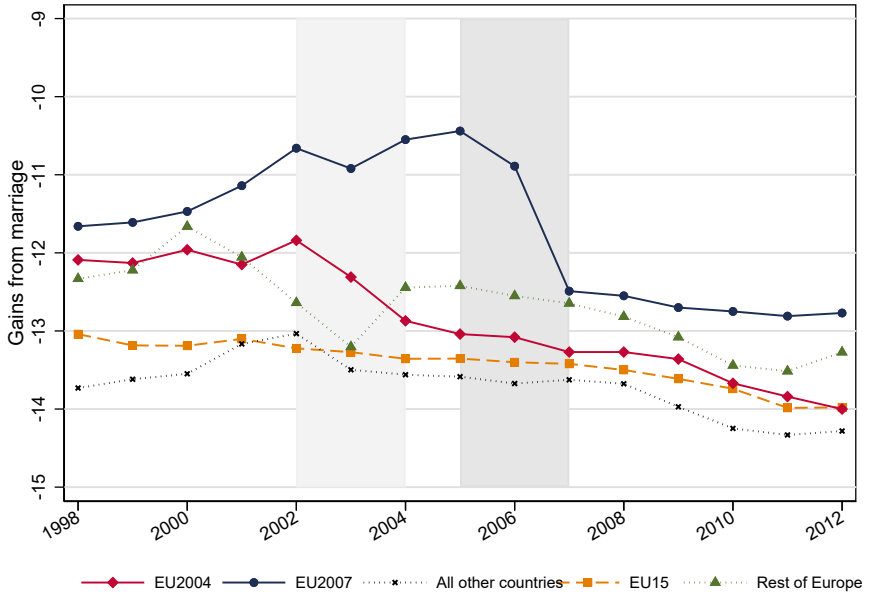
(b) Separation rates, marriages formed since 1998 (index=1 in 2002)



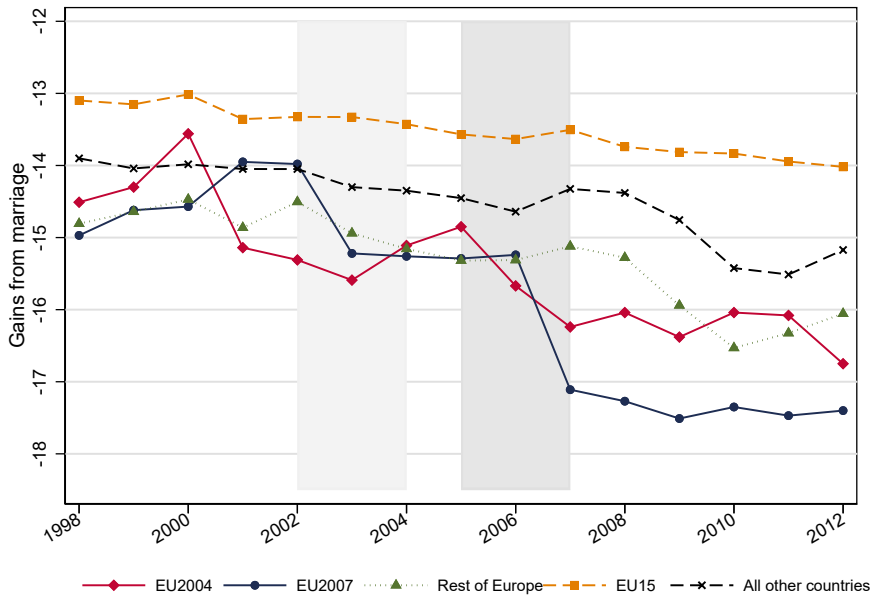
Panel (a) shows the share of marriages in which (at least) one spouse was Italian, by area of origin of the other spouse. For these same marriages, Panel (b) shows the share of separations by area of origin. All series are standardized to 1 in 1998. The classification of countries is reported in Table B1.

Figure 5: Gains from marriage, heterogamous couples formed by natives and immigrants, by area of origin of the foreign spouse

(a) Native husband – foreign wife



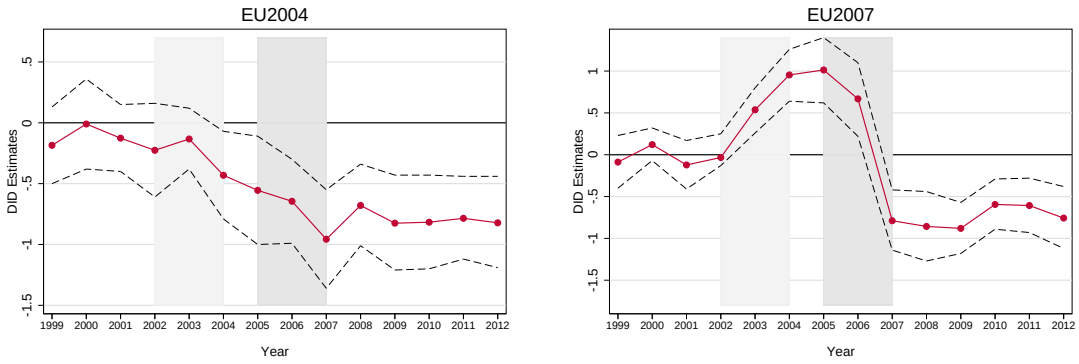
(b) Native wife – foreign husband



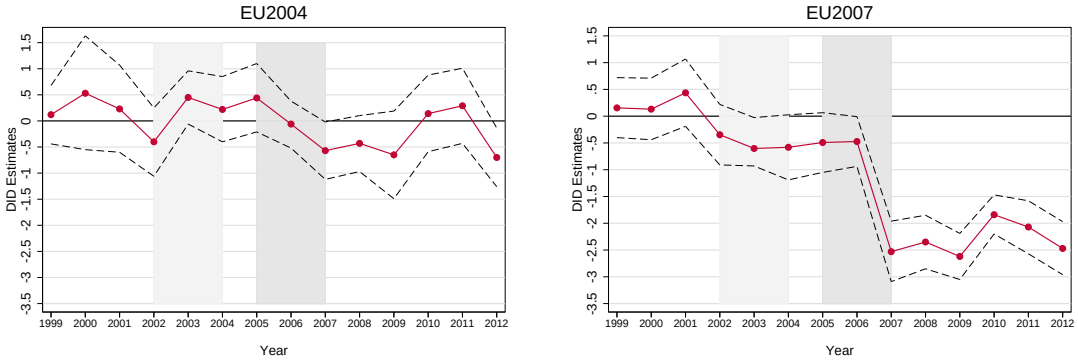
The graphs plot the gains from marriage over the period 1998-2012 by immigrants' area of origin and gender of the foreign spouse. The shaded areas denote the periods between the announcement and implementation of the EU enlargements. Gains from marriage are measured as in equation (1). The classification of countries is reported in Table B1. *Source*: ISTAT, marriage records from vital statistics registries (1998-2012) and individual Census data.

Figure 6: Changes in gains from marriage between natives and new EU citizens, flexible year-specific estimates

(a) Native husband – foreign wife

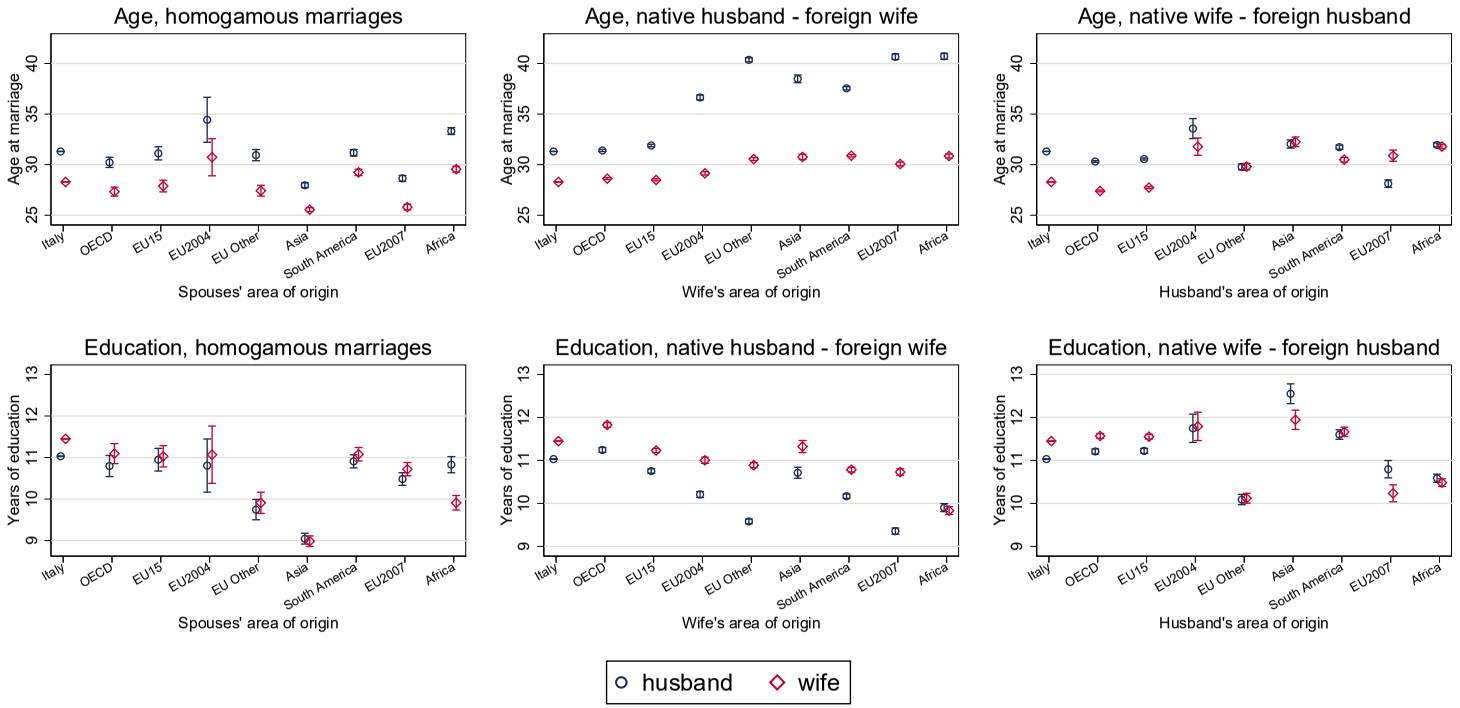


(b) Native wife – foreign husband



This figure shows the estimated effect of admission to the EU on gains from marriage, computed as in equation (1), across cells defined by nationality of the foreign spouse, province, and year. The main explanatory variables are interactions between dummies for intermarriages between natives and new EU citizens (*EU*2004 and *EU*2007) and a full set of year fixed effects. The graphs plot the estimated coefficients and associated confidence intervals, based on standard errors clustered at the province level. Regressions include nationality, province, and year fixed effects, and are weighted by province population.

Figure 7: Spouses' characteristics in homogamous marriages and intermarriages, pre-enlargement period (1998-2002)



The graphs show the average age (top row) and education (bottom row) of husbands and wifes in homogamous marriages (left column) and intermarriages (centre and right columns), by area of origin. 90% confidence intervals are also reported. *Source:* ISTAT, marriage records from vital statistics registries (1998-2002).

Figure 8: Marital surplus related to legal access and cultural affinity and its variation with policy

Before enlargement:

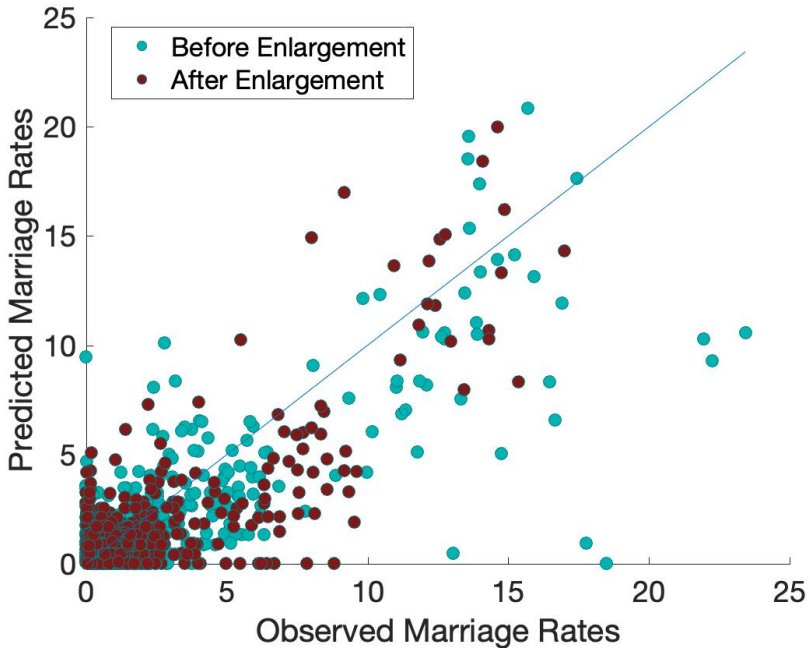
		Women: Italian	Women: New EU	Women: Other
Men	Italian	1	$\bar{\beta}^{\male}\alpha_{I/N}$	$\bar{\beta}^{\male}\alpha_{I/O}$
	New EU	$\bar{\beta}^{\female}\alpha_{N/I}$	$\underline{\beta}\alpha_{N/N}$	$\underline{\beta}\alpha_{N/O}$
	Other	$\bar{\beta}^{\female}\alpha_{O/I}$	$\underline{\beta}\alpha_{O/N}$	$\underline{\beta}\alpha_{O/O}$

After enlargement:

		Women: Italian	Women: New EU	Women: Other
Men	Italian	1	$\alpha_{I/N}$	$\bar{\beta}^{\male}\alpha_{I/O}$
	New EU	$\alpha_{N/I}$	$\alpha_{N/N}$	$\bar{\beta}^{\male}\alpha_{N/O}$
	Other	$\bar{\beta}^{\female}\alpha_{O/I}$	$\bar{\beta}^{\female}\alpha_{O/N}$	$\underline{\beta}\alpha_{O/O}$

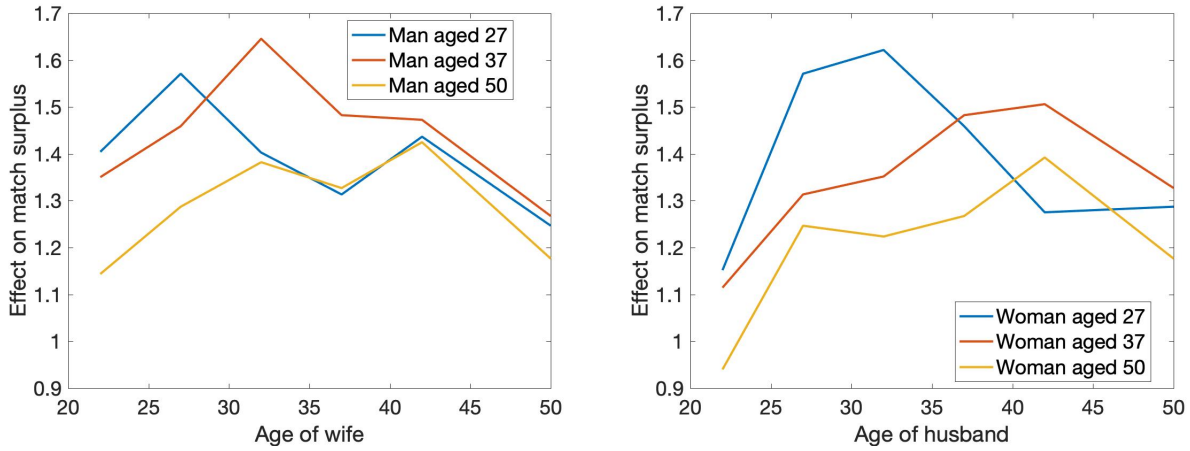
Notes: Each cell shows the marital surplus associated with legal access to work and with culture for spouses of different origins. $\alpha_{x/y}$ are cultural affinity parameters affecting marriage surplus, $\underline{\beta}$ is the penalty for no legal access, $\bar{\beta}^{\female}$ and $\bar{\beta}^{\male}$ parameterize the surplus when either the woman or the man is an EU citizen while the other would not have legal access as single. The surplus has been normalized to one for homogamous marriages between native spouses.
indicates couples who do not have legal access to work in that period.
indicates couples where one of the spouses would not get legal access if single in that period.

Figure 9: Fit of model across marriage markets



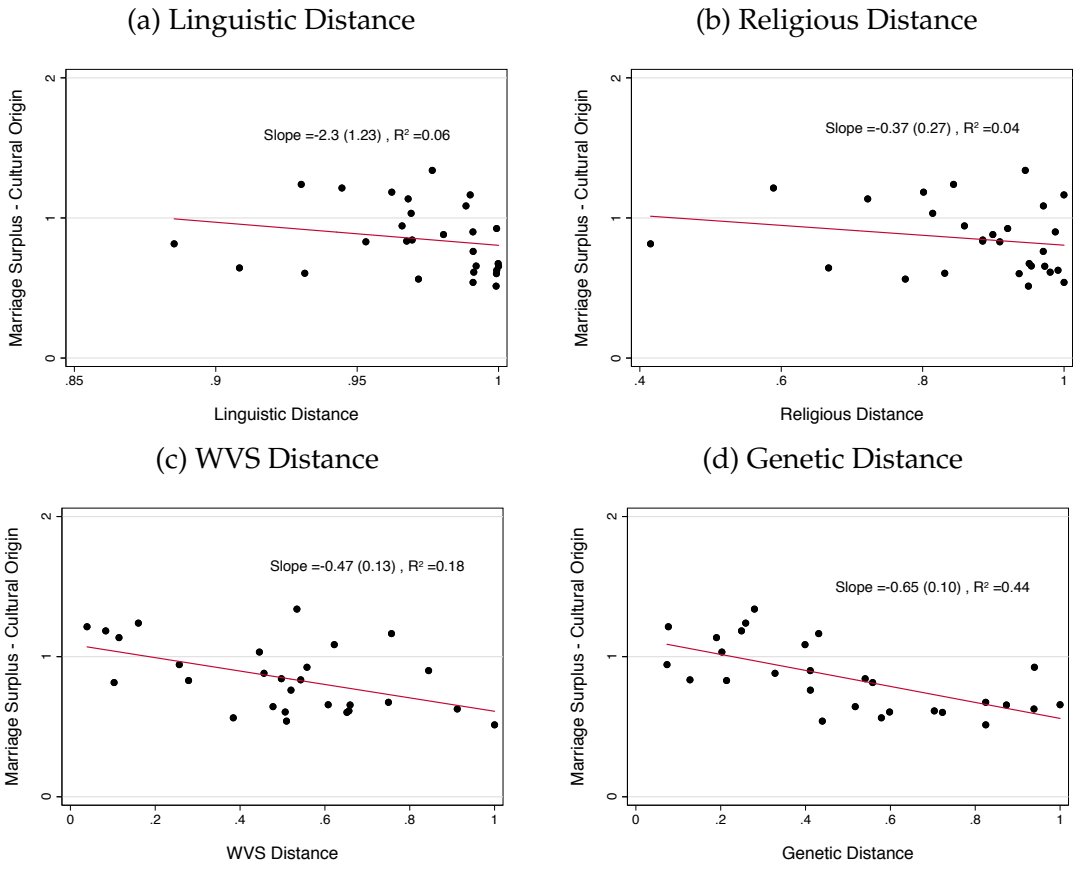
Notes: Each dot represents a particular marriage market, defined by nationality, age, education and province of residence of the spouses. Marriage rates have been multiplied by 1000 for ease of reading.

Figure 10: Marriage surplus, age component by gender



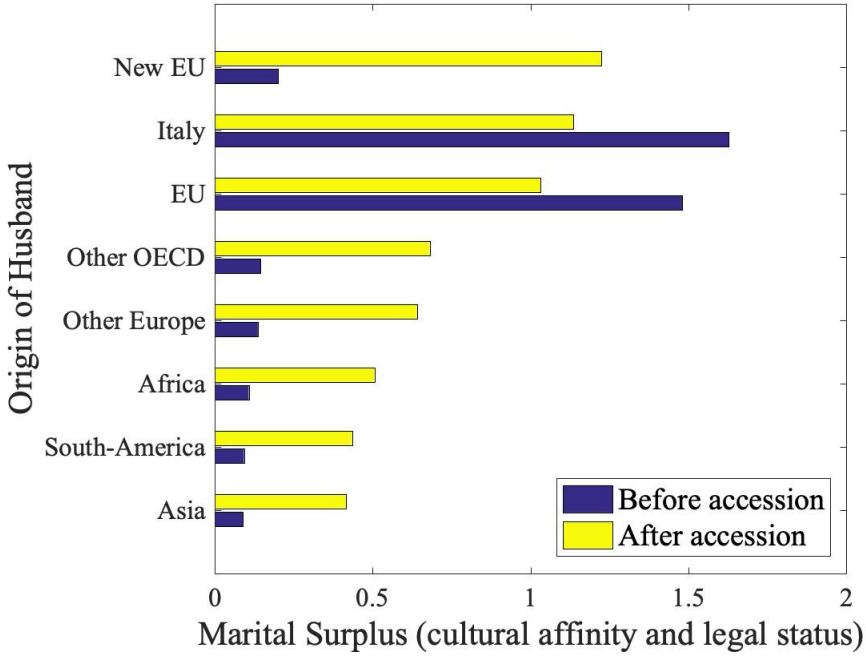
Notes: Parameter estimates of the marital surplus related to age, as defined in equation (7).

Figure 11: Marriage surplus and cultural distance



Note: The Figure plots the relationship between four ex-ante measures of cultural distance from Spolaore and Wacziarg (2016) and the estimated cultural-based component of marriage surplus estimated from the model (see equation 7 and Table 8).

Figure 12: Cultural affinity and legal status, women from New EU countries



Note: The Figure plots the surplus related to cultural affinity and legal status for couples where the wife is a citizen from the New EU countries and the husband is from the list. See equation (7) and Tables 8 and 9 for definitions.

Figure 13: Generalized legal access, change in marriage rates

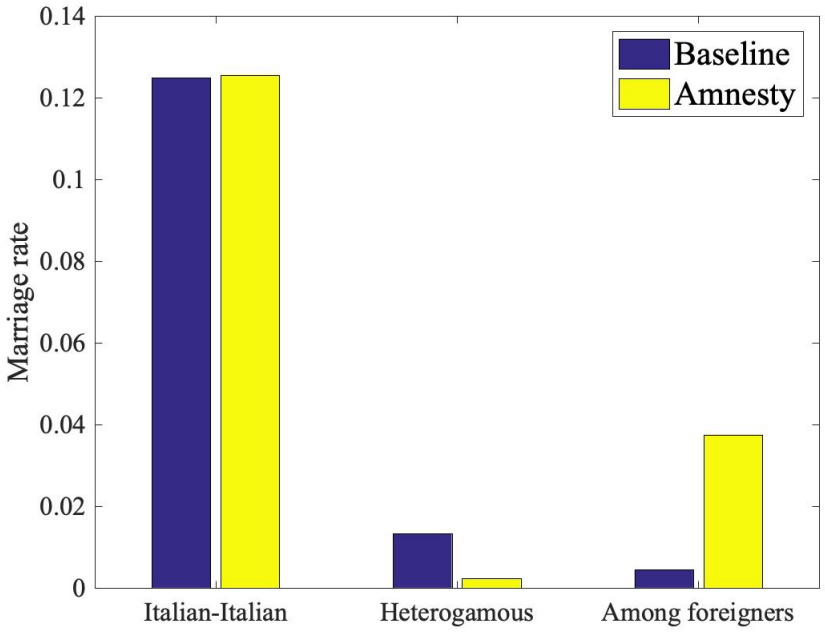


Table 1: Number of observations in the dataset and summary statistics

	Panel A. Marriages			**Panel B. Separations**		
	All	Males	Females	All	Males	Females
Number of observations:						
Natives	6,734,641	3,452,920	3,281,721	381,362	195,934	185,428
Total immigrants	582,941	205,871	377,070	34,762	12,128	22,634
NewEU	100,340	17,702	82,638	5,886	865	5,021
All	7,417,922	3,676,493	3,741,429	422,010	208,927	213,083
Average age:	*at marriage*			*at separation*		
Natives	32.0	33.7	30.2	36.0	37.7	34.2
Total immigrants	32.0	32.4	31.8	34.8	35.8	34.3
NewEU	31.1	30.4	31.3	33.1	33.6	33.0
Average years of education:						
Natives	11.6	11.3	11.9	11.3	11.0	11.6
Total immigrants	11.2	10.9	11.3	10.8	10.7	10.9
NewEU	10.9	10.4	11.1	10.7	10.6	10.8

	Panel C. Singles in 2001			**Panel D. Singles in 2011**		
	All	Males	Females	All	Males	Females
Number of observations:						
Natives	12,996,769	6,917,650	6,079,119	14,317,333	7,587,188	6,730,145
Total immigrants	799,389	428,058	371,331	1,668,957	749,279	919,678
NewEU	60,650	17,125	43,525	377,741	132,732	245,009
All	13,856,808	7,362,833	6,493,975	16,364,031	8,469,199	7,894,832
Average age:						
Natives	31.8	31.1	32.7	34.7	34.3	35.2
Total immigrants	32.0	30.8	33.4	34.9	32.6	36.8
NewEU	32.5	28.6	34.1	34.2	29.2	36.9
Years of education:						
Natives	10.8	10.6	11.0	11.8	11.5	12.1
Total immigrants	10.4	10.0	10.8	10.8	10.3	11.2
NewEU	10.6	9.4	11.0	11.1	10.3	11.5

This table shows the number of individuals included in the marriage and separation registries, and the count of singles in Census years 2001 and 2011; it also reports their average age and years of education.

Table 2: Gains from marriage before and after the EU enlargement, difference-in-differences estimates

Dependent variable:	**Gains from marriage**							
	native male				*native female*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
newEU x *postEU*	-0.673***	-0.650***	-0.767***	-0.719***	-1.645***	-1.524***	-1.452***	-1.402***
	(0.04)	(0.04)	(0.05)	(0.05)	(0.08)	(0.08)	(0.11)	(0.16)
Constant	-8.255***	-12.54***	-7.710***	-13.42***	-14.12***	-13.29***	-14.12***	-14.16***
	(0.09)	(0.00)	(0.14)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	50,349	50,349	50,349	50,349	31,125	31,125	31,125	31,125
R-squared	0.609	0.782	0.645	0.81	0.244	0.854	0.251	0.843
Estimation method	OLS	OLS	WLS	WLS	OLS	OLS	WLS	WLS
Year, prov, country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x year FE	No	Yes	No	Yes	No	Yes	No	Yes
Province x country FE	No	Yes	No	Yes	No	Yes	No	Yes

This table shows the effect of admission to the EU on gains from marriage, estimated using the difference-in-differences specification (2). The sample includes all heterosexual marriages formed between 1998 and 2012 in which at least one of the spouses is Italian. Columns (1)-(4) and (5)-(8) present the results for couples formed by native males and females, respectively. The dependent variable is gains from marriage, computed as in equation (1), across cells defined by province, year, and country of origin of the female spouse (columns 1-4) or male spouse (columns 5-8). The main explanatory variable is the interaction between a dummy for marriages between natives and new EU citizens (*newEU*) and a dummy for the years following their admission to the EU (*postEU*). All regressions include country of origin, province, and year fixed effects; specifications in even columns include, in addition, province $\times$ year and province $\times$ country of origin fixed effects. Regressions in columns (3)-(4) and (7)-(8) are weighted by province population. Standard errors clustered at the province level are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Gains from marriage before and after the EU enlargement, difference-in-differences estimates (robustness to controlling for the number of singles)

Dependent variable:	**Gains from marriage**							
	native male, female from country y				*native female, male from country x*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
newEU x *PostEU*	-0.473***	-0.726***	-0.564***	-0.847***	-0.648***	-1.281***	-0.808***	-0.831***
	(0.06)	(0.07)	(0.08)	(0.09)	(0.14)	(0.13)	(0.13)	(0.11)
$\ln(SingleFemales)$	-0.460***		-0.530***					
	(0.04)		(0.03)					
$\ln(SingleMales_y)$		-0.0706***	0.170***					
		(0.02)	(0.03)					
$\ln(SingleMales)$					-0.722***		-0.820***	
					(0.05)		(0.05)	
$\ln(SingleFemales_x)$						-0.182***	-0.260***	
						(0.04)	(0.04)	
Constant	-0.129	-13.25***	-0.464**	-0.775	1.828***	-13.23***	2.220***	-3.625
	(0.25)	(0.07)	(0.22)	(0.94)	(0.20)	(0.17)	(0.22)	(3.28)
Observations	50,439	44,721	44,721	44,721	31,125	29,539	29,539	29,539
R-Squared	0.817	0.818	0.823	0.836	0.858	0.846	0.86	0.867
Country x Singles	No	No	No	Yes	No	No	No	Yes

This table shows the effect of admission to the EU on gains from marriage, estimated using the difference-in-differences specification (2). The sample includes all heterosexual marriages formed between 1998 and 2012 in which at least one of the spouses is Italian. Columns (1)-(4) and (5)-(8) present the results for couples formed by native males and females, respectively. The dependent variable is gains from marriage, computed as in equation (1), across cells defined by province, year, and country of origin of the female spouse (columns 1-4) or male spouse (columns 5-8). The main explanatory variable is the interaction between a dummy for marriages between natives and new EU citizens (*newEU*) and a dummy for the years following their admission to the EU (*postEU*). All regressions include province $\times$ year and province $\times$ country of origin fixed effects. In addition, columns (1) and (3) include the logarithm of the number of single females in each cell; columns (2) and (3) include the logarithm of the number of single males from country *y*; and column (4) include all these variables interacted with country of origin fixed effects. Analogously, columns (5) and (7) include the logarithm of the number of single males in each cell; columns (6) and (7) include the logarithm of the number of single females from country *x*; and column (8) include all these variables interacted with country of origin fixed effects. Observations are weighted by province population. Standard errors clustered at the province level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Gains from marriage before and after the EU enlargement, difference-in-differences estimates (robustness to alternative ways of computing the gains from marriage)

Dependent variable:	**Gains from marriage**							
	native male, female from country y				*native female, male from country x*			
	no illegals		married+cohabit.		no illegals		married+cohabit.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
newEU × *postEU*	-0.945***	-1.044***	-0.575***	-0.763***	-1.564***	-0.933***	-1.676***	-0.838***
	(0.05)	(0.09)	(0.06)	(0.14)	(0.17)	(0.15)	(0.17)	(0.17)
Constant	-13.32***	-2.589***	-13.50***	-3.475***	-14.06***	-0.821	-14.15***	7.089
	(0.00)	(0.59)	(0.00)	(0.70)	(0.00)	(2.86)	(0.00)	(3.53)
Observations	50,341	44,702	33,885	28,716	31,119	29,534	20,402	19,020
R-squared	0.812	0.839	0.716	0.851	0.837	0.863	0.816	0.876
Nationality x singles	No	Yes	No	Yes	No	Yes	No	Yes

This table shows the effect of admission to the EU on gains from marriage, estimated using different specifications of the difference-in-differences specification (2). The sample includes all heterosexual marriages formed between 1998 and 2012 in which at least one of the spouses is Italian. Columns (1)-(4) and (5)-(8) present the results for couples formed by native males and females, respectively. The dependent variable is gains from marriage, computed as in equation (1), across cells defined by province, year, and country of origin of the female spouse (columns 1-4) or male spouse (columns 5-8). In columns (1)-(2) and (5)-(6) we compute gains from marriage using only information on official foreign residents at the denominator of (1). In columns (3)-(4) and (7)-(8) we add cohabiting, unmarried couples to married couples at the numerator of (1) in order to compute the gains from marriage. The main explanatory variable is the interaction between a dummy for marriages between natives and new EU citizens (*newEU*) and a dummy for the years following their admission to the EU (*postEU*). All regressions include province × year and province × country of origin fixed effects. Specifications in columns (2) and (4) include in addition the logarithm of the number of single females in each cell and the number of single males from country y, both interacted with country of origina fixed effects; columns (6) and (8) include the logarithm of the number of single males in each cell and the number of single females from country x, both interacted with country of origin fixed effects. Observations are weighted by province population. Standard errors clustered at the province level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Hazard rate of separations before and after the EU enlargement, semi-parametric Cox model

Dependent variable:	**Separations**							
	native male, female from country y				*native female, male from country x*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
newEU	0.063***	0.005	0.079	0.068	0.401***	0.351***	0.461***	0.238
	[1.065]	[1.005]	[1.082]	[1.070]	[1.493]	[1.420]	[1.586]	[1.269]
	(0.023)	(0.046)	(0.050)	(0.051)	(0.064)	(0.131)	(0.148)	(0.162)
newEU × *first year after enlargement*	0.227***		0.161***	0.158***	0.149		0.12	0.06
	[1.255]		[1.175]	[1.171]	[1.161]		[1.127]	[1.062]
	(0.050)		(0.054)	(0.054)	(0.144)		(0.154)	(0.158)
newEU × 2 + *years after enlargement*	-0.026		-0.130***	-0.140***	-0.02		-0.165**	-0.335***
	[0.974]		[0.878]	[0.869]	[0.980]		[0.848]	[0.715]
	(0.030)		(0.033)	(0.034)	(0.088)		(0.093)	(0.098)
newEU × *married before enlargement*		0.288***	0.268***	0.276***		0.237***	0.196	0.399***
		[1.334]	[1.307]	[1.318]		[1.267]	[1.217]	[1.490]
		(0.044)	(0.045)	(0.046)		(0.131)	(0.134)	(0.144)
newEU × *married after enlargement*		-0.352***	-0.380***	-0.375***		-0.466***	-0.505***	-0.407***
		[0.703]	[0.684]	[0.687]		[0.628]	[0.604]	[0.666]
		(0.035)	(0.036)	(0.036)		(0.119)	(0.123)	(0.128)
share of residents from country y				0.031				
				[1.031]				
				(0.026)				
share of residents from country x								0.499***
								[1.647]
								(0.030)
Observations	11,182,250	11,182,250	11,182,250	11,176,994	10,683,818	10,683,818	10,683,818	10,681,099
Year FE	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Stratified by year of marriage	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Pool at risk (marriages)	3,426,235	3,426,235	3,426,235	3,424,486	3,255,923	3,255,923	3,255,923	3,255,087
Separations	195,879	195,879	195,879	195,777	185,356	185,356	185,356	185,300

This table shows the effect of admission to the EU on hazard rate of separation, modeled using a semi-parametric Cox model. The sample includes all marriages formed between 1998 and 2012 in which at least one of the spouses is Italian. Columns (1)-(4) present the results for couples in which the husband is native, whereas columns (5)-(8) present the results for couples in which the wife is native. The main explanatory variables of interest are a dummy for intermarriages between natives and new EU citizens (*newEU*) and its interactions with dummies for the first year after the enlargement; the subsequent years; couples formed before the enlargement; and couples formed after the enlargement. The specifications in columns (4) and (8) also include on the right-hand side the share of population in the province from the same country of origin as the foreign spouse. The baseline hazard rate is stratified by native-native vs. native-immigrant couples and province in all specifications, and by year of marriage in columns (2)-(4) and (6)-(8). Calendar year fixed effects are included in columns (1), (3)-(4), (5), and (7)-(8). Hazard ratios are reported in brackets and robust standard errors clustered by marriage are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Divorce regressions, observed and predicted coefficients

	Before enlargement		During enlargement	
	Observed	Predicted	Observed	Predicted
Italian-New EU	0.0071 (0.0087)	-0.0056	0.0098 (0.0101)	0.0113
Italian-Others	0.0018 (0.0261)	-0.0052	0.0039 (0.0284)	-0.0055
Non Italian - Non Italian	-0.0047 (0.0685)	0.0015	-0.0036 (0.0391)	-0.0058
Woman >10 year younger	0.0011 (0.0189)	-0.0015	0.0019 (0.017)	0
Woman >10 year older	0.0029 (0.4)	-0.0055	0.0099 (0.438)	-0.0048
Same education	0.0001 (0.0722)	-0.0024	0.0011 (0.0504)	-0.0023
Constant	0.0122 (0.0104)	0.0081	0.0108 (0.0073)	0.0081

The table displays the auxiliary coefficients in a linear regression of an indicator variable equal to one if a couple separate within 3-4 years after marriages. Two separate regressions are run for couples married and followed before the enlargements and for couples married before and followed through the enlargements. The regression is defined in equation (9). Standard errors are displayed in parenthesis.

Table 7: Estimated parameters, marriage surplus linked to education and home ownership

		Women					
		Education					
		Low	High				
Men	Low	1.09	0.901	Housing, natives			
	High	0.96	1	No	Yes		
			No	1	1.58	Housing, non natives	
			Yes	1.76	1.76	No	Yes
					No	1	1.1
					Yes	1.14	2.0

Notes: The table displays the part of the surplus associated with education and housing, as defined in equation (6).

Table 8: Estimated parameters, marriage surplus by area of origin

		Women							
		Italian	EU15	New EU	Other Europe	Africa	Asia	South America	OECD
Men	Italian	1	1.2	1.1	0.94	0.92	0.76	0.82	1.2
	EU	1.2	1.1	1	0.83	0.51	0.9	0.64	1.2
	New EU	1.1	1	1.2	0.83	0.66	0.54	0.56	0.88
	Oth. Europe	0.94	0.83	0.83	6.8	0.67	1.1	0.84	1.3
	Africa	0.92	0.51	0.66	0.67	6.1	0.66	0.6	0.63
	Asia	0.76	0.9	0.54	1.1	0.66	4.6	0.61	1.2
	S America	0.82	0.64	0.56	0.84	0.6	0.61	6.6	0.6
	OECD	1.2	1.2	0.88	1.3	0.63	1.2	0.6	9.7

Notes: The table displays the part of the surplus associated with country of origin, as defined in equation (6).

Table 9: Estimated parameters: surplus linked to legal access, outside value and shocks

Parameter	Value
Panel A: legal access	
No legal access ($\underline{\beta}$), North	0.165
No legal access ($\underline{\beta}$), South	0.265
Legal access, native husband - foreign wife ($\bar{\beta}^{\male}$), North	1.43
Legal access, native wife - foreign husband ($\bar{\beta}^{\female}$), North	0.772
Legal access, native husband - foreign wife ($\bar{\beta}^{\male}$), South	1.42
Legal access, native wife - foreign husband ($\bar{\beta}^{\female}$), South	1.41
Panel B: outside value ω	
Before enlargement, North	6.08
Before enlargement, South	5.76
After enlargement, North	6.06
After enlargement, South	5.74
Panel C: love component and separation	
Standard error of "love" component (σ_ε)	0.27
Innovation of "love" component (σ_η), native homogamous marriages	0.03
Innovation of "love" component (σ_η), mixed of foreign marriages	0.08
Fraction non-divorcers (δ), natives	0.82
Fraction non-divorcers (δ), foreigners	0.97

Notes: The table displays in Panel A the part of the surplus associated with legal access obtained through marriage, as defined in equation (6) (see also Figure 8 for a description). Panel B displays the value of remaining single (see Section 4.1.1 for a description). Panel C displays the parameters related to the couple specific "love" shock (see Section 4.1.3).

APPENDIX

A Imputation of the total number of unmatched singles

In this section we describe the procedure we use to impute the total number of unmatched singles, both among official and unofficial immigrants, at the denominator of the gains from marriage in equation (1). Unmatched singles are defined throughout as individuals aged between 18 and 60 that are not married.

A.1 Official immigrants

We obtained from the Italian Census the exact number of singles by age, gender, nationality, and province in years 1991, 2001, and 2011, and combined these data with migration inflows at the same level of aggregation during the period 1995-2012 to estimate the stock of immigrants in each cell. Let S_{ct} denote the number of migrant singles in year $t = 1991, 2001, 2011$, for cell c (age, nationality, gender, province); and M_{ct} denote the total number of migrants for cell c in years from 1995 to 2012 (irrespective of marital status). Moreover, let $\Delta S_{c,(t-s)}$ and $\Delta M_{c,(t-s)}$ the changes in S and M between any two years, t and s.

Consider first the period 2001-2011. We assume that the yearly change in the number of singles in each cell between any two consecutive years, as a share of the total change over the period 2001-2011, equals the corresponding change for total migrants:

$$\frac{\Delta M_{c,(t-(t-1))}}{\Delta M_{c,(2011-2001)}} = \frac{\Delta S_{c,(t-(t-1))}}{\Delta S_{c,(2011-2001)}}.$$

From the above equation, we recover the number of single migrants over time starting from $S_{c,2001}$ and adding up the yearly variation, $\Delta S_{c,(t-(t-1))} = \Delta M_{c,(t-(t-1))} * \Delta S_{c,(2011-2001)} / \Delta M_{c,(2011-$
We implement the same procedure for the earlier period exploiting information on total migrants and singles in Census years 1991 and 2001 as well as migration inflows over the period 1995-2001.

A.2 Unofficial immigrants

As discussed in Section 3.2.2, neglecting unofficial migrants would bias upward the estimated gains from marriage in equation (1). Most importantly, the bias would be asymmetric between new EU citizens and other immigrants, before and after the EU enlargements. We thus adjust the number of immigrant singles for the presence of irregulars using information from three sources: (i) number of singles and non-singles among regular immigrants by cell (as defined, e.g., by province and nationality) in years 1991, 2001, and 2011, from the national Census; (ii) number of singles and non-singles among irregular immigrants by cell in year 2002, from applications to a generalized amnesty; (iii) total (estimated) number of irregular immigrants.

We next discuss our method for imputing irregular immigrants, and the data on irregular immigrants that we use to implement such method.

A.2.1 Methodology

Let S^R_{c,t^*} and S^I_{c,t^*} be the number of singles among regular and irregular immigrants, respectively, by cell c in a benchmark year t^*. Therefore, the total number of singles among immigrants in year t^* is $S_{c,t^*} = S^R_{c,t^*} + S^I_{c,t^*}$. Moreover, let M^R_t and M^I_t be the number of regular and irregular immigrants (including both singles and non-singles) in year t, so $M_t = M^R_t + M^I_t$ is the number of total immigrants in the same year. We impute the (unobserved) number of irregular immigrant singles outside the benchmark year, $S^I_{c,t}$ $\forall t \neq t^*$, based on knowledge of $S^R_{c,t}$, M^R_t, and M^I_t, under the following assumptions.
Assumption 1: The share of irregulars among singles is proportional to the share of irregulars among total immigrants:

$$\frac{S^I_t}{S_t} = \gamma \frac{M^I_t}{M_t}, \tag{11}$$

where $S^I_t = \sum_c S^I_{c,t}$ and $S_t = \sum_c S_{c,t}$, and γ is a constant.
Assumption 2: The ratio of irregulars to regulars among singles follows common trends across cells:

$$\frac{S^I_{c,t}}{S^R_{c,t}} = \alpha_c \beta_t, \tag{12}$$

where α_c and $beta_t$ are cell-specific and sector-specific constants.

In the next section, we provide indirect evidence consistent with these two assump-

tions. Under such assumptions, we impute the total number of immigrant singles (including both regular and irregular immigrants) in three steps. First, we use the total number of regular and irregular immigrants in the benchmark year $t^* = 2002$ – the latter being estimated from amnesty applications – to compute γ in equation (11):

$$\gamma = \frac{S^I_{t^*} / S_{t^*}}{M^I_{t^*} / M_{t^*}}$$

Second, we aggregate individual-level data on regular and irregular immigrants by cell and normalize $\beta_{t^*} = 1$ to obtain

$$\alpha_c = \frac{S^I_{c,t^*}}{S^R_{c,t^*}}, \ \forall c.$$

Third, we combine (11) and (12) to compute the β's for all other years:

$$\beta_t = \frac{\gamma M^I_t / M_t}{1 - \gamma M^I_t / M_t} \frac{\sum_c S^R_{c,t}}{\sum_c \alpha_c S^R_{c,t}}, \ \forall t.$$

All terms on the right-hand side of the last equations are either data or parameters estimated in previous steps.

Finally, the total number of immigrants (including both regular and irregular immigrants) in cell c and year t equals

$$S_{c,t} = (1 + \alpha_c \beta_t) S^R_{c,t}$$

.

A.2.2 Data

In general, it is difficult to precisely quantify the number of irregular immigrants, as the latter typically hide from administrative authorities. Generalized amnesties of irregular immigrants represent an important exception, as in these occasions applicants can obtain a residence permits under relatively mild conditions. The number of applications thus provides a fairly accurate estimate of the number of irregular immigrants.

In Italy, five such amnesties were granted in 1986, 1990, 1995, 1998 and 2002. We obtained individual-level data on all applicants to the 2002 amnesty, which granted res-

idence and work permits to about 700 thousand immigrants – the largest number ever legalized in Italy.[28] We use the total number of applications to estimate the total number of irregular immigrants in the benchmark year $t^* = 2002$, $M^I_{t^*}$.

We estimate S^I_t and S^I_{c,t^*} using additional information on gender, marital status, age, province of residence, country of origin, hours of work, and monthly wage of applicants, also contained in the amnesty files.[29] Since the cells in our model (i.e., $\bar{x}$ and $\bar{y}$ in Section 4) are defined also by education, but the amnesty files do not report applicants' educational attainment, we compute the predicted probability of having a college degree conditional on other individual characteristics. Specifically, we estimate such probability using data from a representative survey of irregular immigrants conducted by the Institute for Multi-ethnic Studies (ISMU).

ISMU is a research institute collecting data on immigrants in Italy, focusing in particular on undocumented migrants. Starting in 2001, it has been conducting an annual survey on a representative sample of about 8,000 regular and irregular migrants residing in the Italian region of Lombardy. This survey is specifically designed to elicit truthful reporting of legal status. In addition, the survey contains detailed information on other individual characteristics, including gender, age, education, country of origin, marital status, employment, wage and province of residence.[30] We thus regress a dummy for college education on gender, age, age squared, wage class and country of origin fixed effects, and we use the estimated coefficients to impute college education across amnesty applicants.

ISMU also provides an estimate of the total number of irregular immigrants in each year since 1991 (Figure C.1), which we use to measure M^I_t, $\forall t \neq t^*$.

[28] We thank Francesco Fasani for kindly sharing these data; see Fasani (2009) for a detailed description of generalized amnesties in Italy.

[29] Table C.1 provides some descriptive statistics, separately for male and female applicants – 377 and 318 thousand individuals, respectively.

[30] Table C.2 provides descriptive statistics on individual characteristics for regular and irregular migrants. These data have been used, among others, by Dustmann et al. (2017) and Guriev et al. (2016) Additional information is available through the website www.ismu.org.

B Appendix: additional figures and tables

Table B1: Classification of countries by geographic area of origin

EU15	Austria, Belgium, Denmark, Finland, Germany, Ireland, Luxembourg, Netherlands, United Kingdom, Sweden, France, Greece, Portugal, Spain.
EU2004	Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovakia, Slovenia.
EU2007	Bulgaria, Romania.
EU Other	Andorra, Belarus, Isle of Man, Liechtenstein, Norway, Monaco, Republic of Moldova, Russian Federation, San Marino, Ukraine, Vatican City State, Albania, Bosnia and Herzegovina, Croatia, Iceland, Kosovo, Macedonia (FYROM), Serbia, Montenegro, Turkey.
Africa	Algeria, Egypt, Libyan Arab Jamahiriya, Marocco, Tunisia, 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.
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, 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.
OECD Other	Australia, Canada, Japan, Republic of Korea, New Zealand, Norway, Switzerland, United States.

Table B2: Spouses' characteristics, before and after the EU enlargement

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**difference in age**		**years of education**		**educ(male)>educ(female)**	
female from NewEU x PostEU	-0.487***	-0.296***	0.058	0.030	0.028**	0.025**
	(0.103)	(0.104)	(0.066)	(0.074)	(0.011)	(0.012)
male from NewEU x PostEU	0.531***	1.554***	-0.078	-0.527***	-0.031**	-0.106***
	(0.203)	(0.441)	(0.085)	(0.145)	(0.014)	(0.025)
Constant	3.331***	3.334***	-0.396***	-0.404***	-0.073***	-0.075***
	(0.015)	(0.015)	(0.006)	(0.007)	(0.001)	(0.001)
Observations	3,268,228	3,201,479	3,268,228	3,201,479	3,268,228	3,201,479
all marriages	x		x		x	
at least 1 native spouse		x		x		x
R-squared	0.072	0.072	0.004	0.004	0.002	0.002

This table shows the effect of the EU enlargement on differences in age and education between husband and wife. The dependent variable in columns (1) and (2) is the difference in age; in columns (3) and (4), it is the difference in years of education; and in columns (5) and (6) it is an index equal to 1 if the husband is more educated than the wife, 0 if they have the same level of education, and -1 if the wife is more educated than the husband. The equation is estimated at the marriage level and the sample includes all marriages formed during the period 1998-2012 (odd columns) or only marriages in which at least one of the spouses is a native (even columns). The explanatory variables of interest are dummies equal to 1 when the husband or the wife is a citizen from new EU countries, interacted with a dummy equal to 1 for the years after the EU enlargement. Fixed effects for spouses' nationality pairs are included in all specifications. Robust standard errors clustered by nationality pair-year cell are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table B3: Gains from marriage before and after the EU enlargement, difference-in-differences estimates allowing for announcement effects and for substitution across nationalities

Dep Var:	**Gains from marriage**							
	Native husband, foreign wife				*Native wife, foreign husband*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EU2004 × Post2003	-0.003	0.045	0.007	0.044	0.194	0.019	0.440	0.312
	(0.073)	(0.080)	(0.077)	(0.093)	(0.187)	(0.199)	(0.211)	(0.197)
EU2004 × Post2004	-0.228***	-0.237***	-0.359***	-0.355***	-0.074	-0.181	-0.119	-0.352
	(0.073)	(0.084)	(0.127)	(0.145)	(0.186)	(0.210)	(0.234)	(0.199) *
EU2004 × Post2006	-0.170**	-0.111	-0.154	-0.095*	-0.255	-0.285	-0.398	-0.357
	(0.068)	(0.069)	(0.098)	(0.103)	(0.193)	(0.226)	(0.253)	(0.208) *
EU2004 × Post2007	-0.099	-0.169***	-0.168**	-0.204***	-0.372**	-0.154	-0.259	-0.004
	(0.065)	(0.067)	(0.083)	(0.074)	(0.182)	(0.200)	(0.231)	(0.303)
EU2007 × Post2003	0.593***	0.706***	0.564***	0.703***	-0.673***	-0.703***	-0.666***	-0.674***
	(0.091)	(0.095)	(0.099)	(0.093)	(0.128)	(0.141)	(0.136)	(0.163)
EU2007 × Post2004	0.494***	0.510***	0.446***	0.482***	0.167	0.233	0.067	0.148
	(0.082)	(0.089)	(0.091)	(0.099)	(0.139)	(0.151)	(0.175)	(0.192)
EU2007 × Post2006	-0.405***	-0.439***	-0.315***	-0.346***	-0.046	-0.025	0.062	0.080
	(0.075)	(0.084)	(0.098)	(0.104)	(0.145)	(0.170)	(0.154)	(0.184)
EU2007 × Post2007	-1.302***	-1.292***	-1.415***	-1.417***	-1.725***	-1.819***	-1.850***	-1.951***
	(0.075)	(0.080)	(0.104)	(0.112)	(0.134)	(0.159)	(0.201)	(0.237)
Observations	48,933	48,933	48,933	48,933	29,698	29,698	29,698	29,698
R-squared	0.474	0.706	0.552	0.760	0.497	0.762	0.541	0.775
Estimation method	OLS	OLS	WLS	WLS	OLS	OLS	WLS	WLS
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nationality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province × Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Province × Nationality FE	No	Yes	No	Yes	No	Yes	No	Yes

This table shows the estimated effect of admission to the EU on gains from marriage, computed as in equation (1), across cells defined by nationality of the foreign spouse, province, and year. The main explanatory variables are interactions between dummies for marriages between natives and new EU citizens (*EU*2004 and *EU*2007) and dummies for the periods after the announcement (*P*2003 and *P*2006) and implementation of each enlargement (*P*2004 and *P*2007). All regressions include nationality, province, and year fixed effects; specifications in even columns include, in addition, province × year and province × nationality fixed effects. Columns (1)-(4) present the results for couples in which the husband is native, whereas columns (5)-(8) present the results for couples in which the wife is native. Regressions in columns (3)-(4) and (7)-(8) are weighted by province population. Standard errors clustered at the province level are reported in parentheses (99 provinces in total). Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

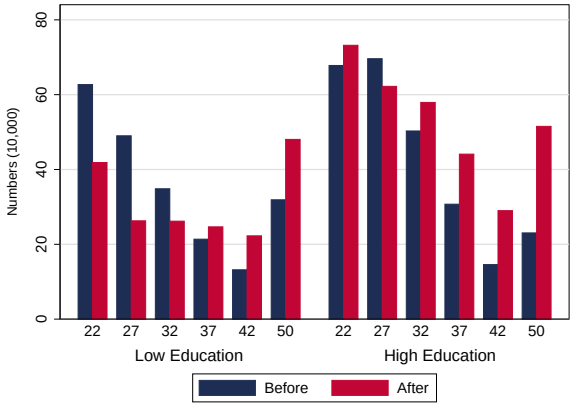
Table B4: Internet diffusion and intermarriages

Dependent variable:	*native male, female from country y*				*native female, male from country x*			
	log n. marriages		**gains from marriage**		**log n. marriages**		**gains from marriage**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln(SingleFemales_y)$	0.211***	0.258***						
	(0.010)	(0.0127)						
$\ln(SingleFemales_y) \times Internet$		-0.003***						
		(0.0004)						
$\ln(SingleMales_x)$					0.095***	0.172***		
					(0.011)	(0.016)		
$\ln(SingleMales_x) \times Internet$						-0.005***		
						(0.001)		
$newEU \times PostEU$			-0.793***	-0.448**			-1.381***	-1.759***
			(0.055)	(0.175)			(0.143)	(0.257)
$newEU \times PostEU \times Internet$				-0.021**				0.023
				(0.010)				(0.021)
Constant	0.434***	0.486***	-13.48***	-13.51***	0.650***	0.678***	-14.26***	-14.27***
	(0.045)	(0.043)	(0.003)	(0.004)	(0.052)	(0.053)	(0.005)	(0.005)
Observations	39,313	39,313	39,313	39,313	23,793	23,793	23,793	23,793
R-squared	0.924	0.924	0.814	0.814	0.928	0.928	0.852	0.852

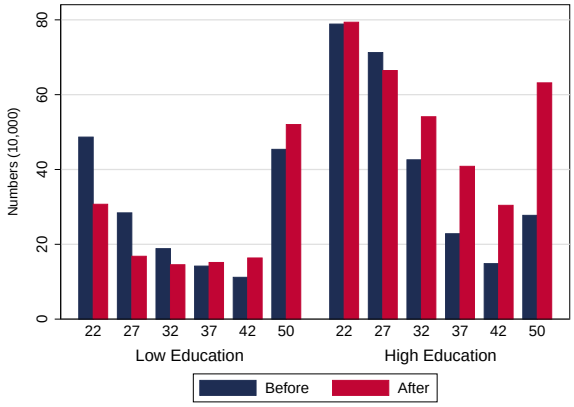
This table shows the effect of Internet diffusion on the formation of marriages between natives and foreigners, across cells defined by nationality of the foreign spouse, province, and year. The dependent variable in columns (1)-(2) and (5)-(6) is the log number of marriages in each cell, whereas in columns (3)-(4) and (7)-(8) the dependent variable is gains from marriage, computed as in equation (1). The variable *Internet* is the share of population with access to broadband internet connection in each region and year, as collected by the ISTAT survey *Aspects of Daily Life Survey*. The other explanatory variables are defined as in Table 3. All regressions include province $\times$ year and province $\times$ nationality fixed effects and they are weighted by province population. Standard errors clustered at the province level are reported in parentheses (99 provinces in total). Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Figure B1: Distribution of population vectors, by age and area of origin

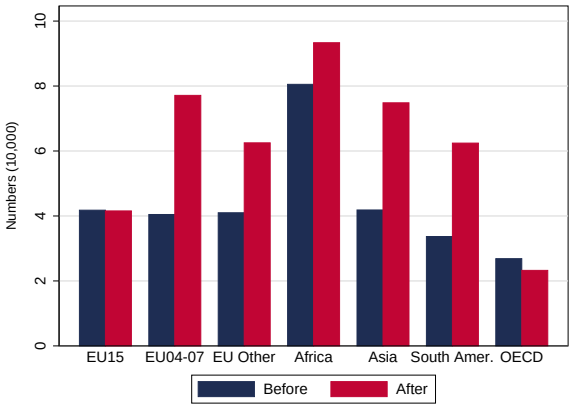
(a) Men by Age and Education



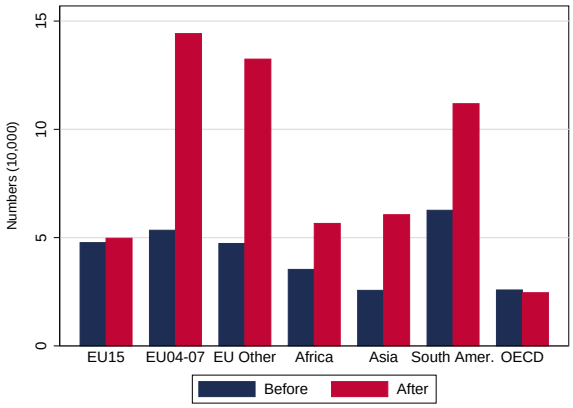
(b) Women by Age and Education



(c) Men by area of origin

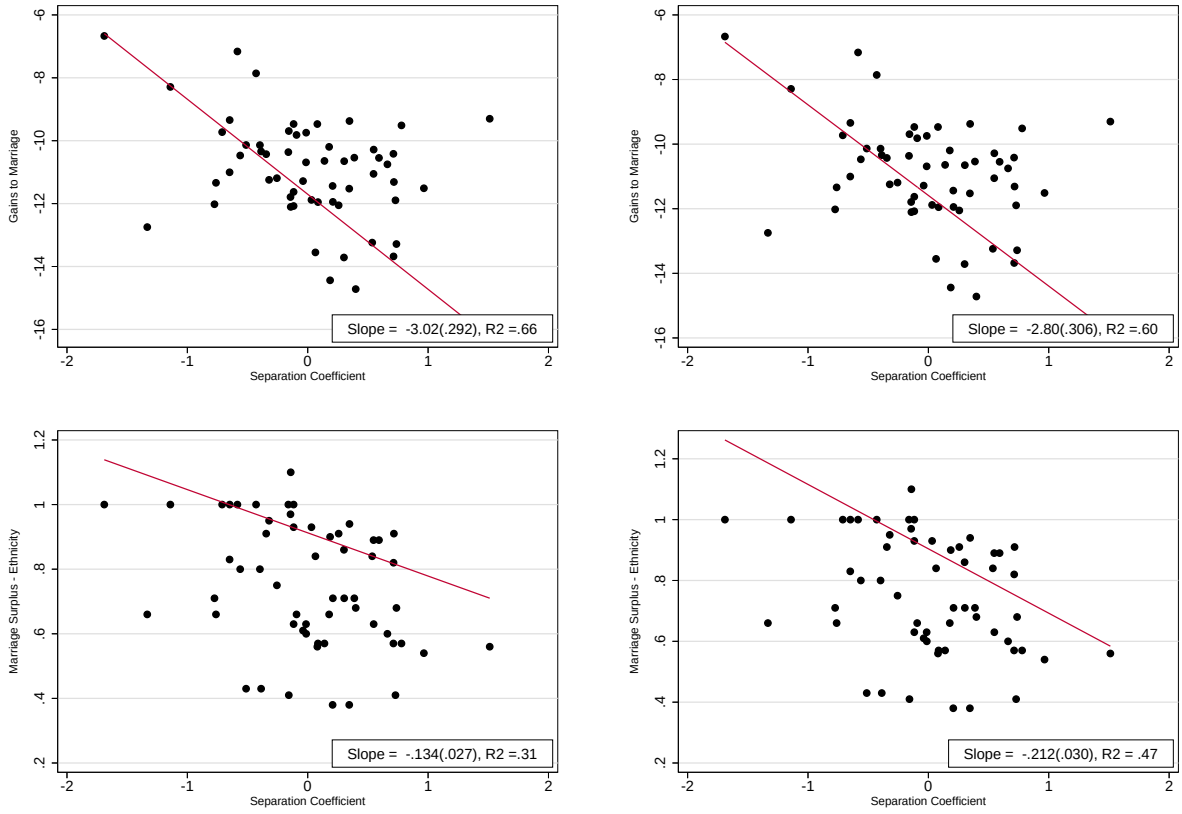


(d) Women by area of origin



Source: Marriage (1998-2002 and 2007-2012) and individual Census (2001, 2011) data, Italy. The Figures report the distributions of men and women population vectors by age and education and area of origin, separately.

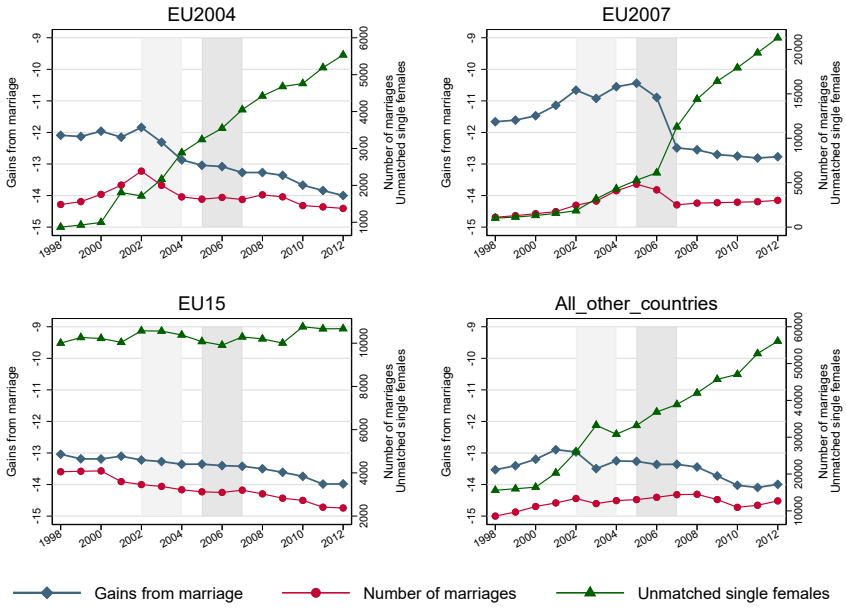
Figure B2: Gains from Marriage/Model Marital Surplus and Probability of Separation



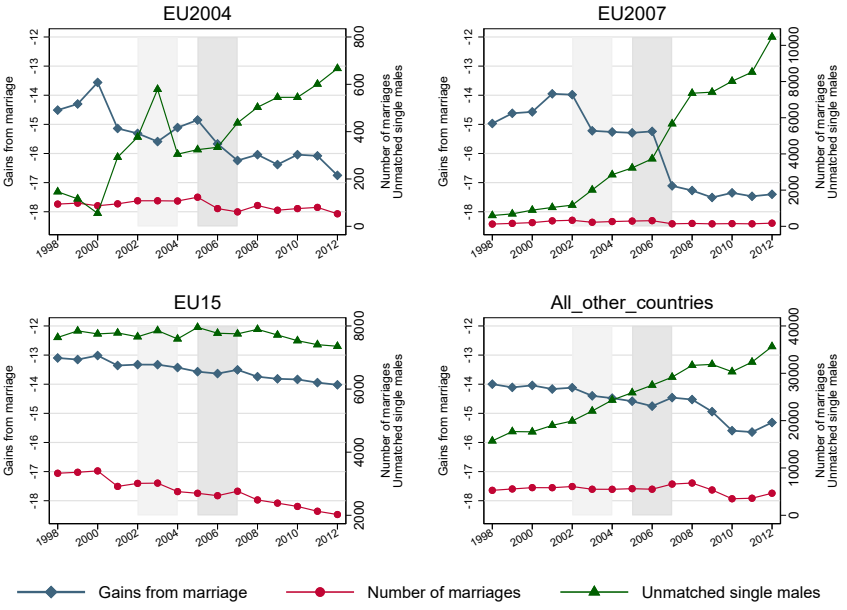
Note: The Figure plots the relationship between the gains from marriage by area of origin and the separation coefficients estimated from a Cox duration model by match. Marriages formed after the two enlargements, from 2007 to 2012.

Figure B3: Unmatched singles, marriages, and gains from marriage by area of origin, 1998-2012

(a) Native husband – foreign wife

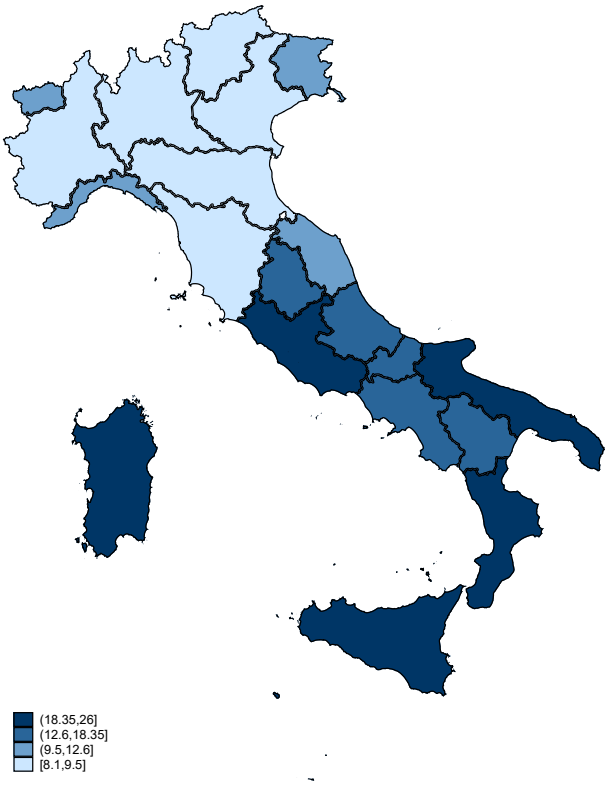


(b) Native wife – foreign husband



The graphs plot the number of unmatched singles, the number of marriages, and the gains from marriage over the period 1998-2012 by area of origin of the foreign spouse. The shaded areas denote the periods between the announcement and implementation of the EU enlargements. Gains from marriage are measured as in equation (1). The classification of countries is reported in Table B1. *Source*: ISTAT, marriage records from vital statistics registries (1998-2012) and individual Census data.

Figure B4: Size of the shadow economy across Italian regions



Source: Istat Shadow Economy Dossier (2010), Italy. The map shows the size of the shadow economy across Italian regions in 2002, as measured by the estimated fraction of irregular workers over total workers .

Table C.1: Characteristics of applicants to the generalized amnesty of 2002

	Males		Females	
	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics				
Age	30.04	7.62	35.16	9.97
Married	0.424	0.490	0.546	0.500
Geographic area of origin				
EU2004	0.024	0.150	0.093	0.290
EU2007	0.219	0.410	0.215	0.410
EU Other	0.229	0.420	0.414	0.490
Africa	0.271	0.440	0.053	0.220
Asia	0.188	0.390	0.077	0.270
South America	0.068	0.250	0.146	0.350
OECD	0.001	0.030	0.001	0.040
Area of residence in Italy				
North	0.549	0.500	0.483	0.500
Center	0.271	0.440	0.314	0.460
South	0.180	0.380	0.203	0.400
Observations	376,840		317,964	
Total Observations	694,804			

Figure C.1: Regular and irregular immigrants, 1991-2013

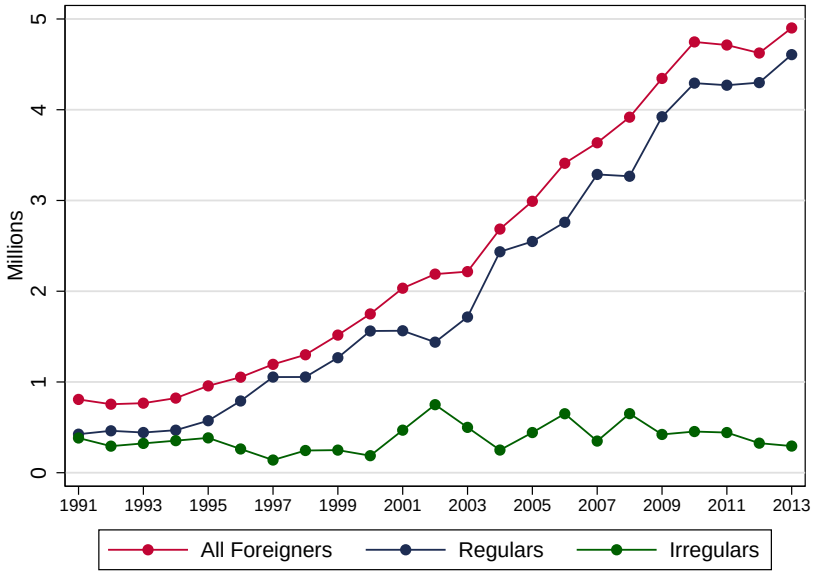


Table C.2: ISMU survey, summary statistics for regular and irregular immigrants

	Regular migrants				*Irregular migrants*			
	Males		Females		Males		Females	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics								
Age	35.56	8.33	35.36	8.70	30.18	7.03	33.83	9.22
Married	0.625	0.480	0.633	0.480	0.286	0.450	0.389	0.490
Children (#)	1.270	1.400	1.360	1.210	0.596	1.130	1.100	1.240
High Education	0.519	0.500	0.608	0.490	0.452	0.500	0.596	0.490
Employed	0.886	0.320	0.644	0.480	0.731	0.440	0.777	0.420
Legal Job	0.813	0.390	0.550	0.500	0.059	0.240	0.067	0.250
Illegal Job	0.073	0.260	0.094	0.290	0.672	0.470	0.710	0.450
Income (Euro)	998.94	697.50	533.71	553.06	553.29	547.53	580.91	519.93
Geographic area of origin								
EU2004	0.004	0.060	0.021	0.140	0.002	0.040	0.005	0.070
EU2007	0.040	0.200	0.077	0.270	0.052	0.220	0.080	0.270
EU Other	0.136	0.340	0.209	0.410	0.135	0.340	0.342	0.470
Africa	0.525	0.500	0.321	0.470	0.561	0.500	0.166	0.370
Asia	0.223	0.420	0.169	0.380	0.141	0.350	0.084	0.280
South America	0.072	0.260	0.203	0.400	0.109	0.310	0.324	0.470
OECD	0.000	0.010	0.000	0.020	0.000	0.002	0.000	0.000
Observations	51,181		38,897		7,835		3,822	
Total Observations	90,078				11,657			

Note: ISMU (2001-2016) data, Lombardy. Individual level data.