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

Jérôme Adda† Paolo Pinotti‡ Giulia Tura§

April 18, 2024

Abstract

We analyse the contribution of legal status incentives on the marriage choices of natives and migrants. Access to legal status reduces by 40 percent the probability of immigrants intermarrying with natives, and increases by 20 percent the hazard rate of separation for intermarriages. We develop and estimate a multidimensional equilibrium model of marriage, fertility, and separation, where individuals match on observed and unobserved characteristics. Allowing for trade-offs between cultural distance, legal status, and other socio-economic spousal characteristics, we quantify the role of legal status and the strength of cultural preferences and evaluate the welfare consequences of granting legal status to immigrants.

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

*The data on marriages and separations used in this paper have been accessed through the Laboratory for the Analysis of Elementary Data (ADELE) at ISTAT, in compliance with the laws on the protection of statistical confidentiality and of personal data. We are solely responsible for the results and the opinions expressed in this paper, which do not constitute official statistics. We 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. Marta Morando provided excellent research assistance.

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

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

§Dept. of Economics, Management and Statistics, University of Milano-Bicocca. email: giulia.tura@unimib.it.

1 Introduction

The increasing migration trend has led to challenges for receiving countries in assimilating foreign-born individuals. These challenges are both economic and cultural and are crucially interdependent. For instance, in many countries, immigrants who marry natives gain legal access to work and residency. Hence, policies targeting one assimilation dimension may spill over and influence the other margin. There is, therefore, a need to better understand how these two margins are intertwined to design effective policies.

This paper studies the effect of granting legal access to migrants on marriage markets and how it affects marital stability and fertility patterns. Inter-marrying involves several trade-offs. The foreign spouse has to weigh the benefit of getting access to the labor market against costs, as natives may be less desirable spouses than co-nationals due to cultural distance or different attributes. Natives face similar trade-offs as foreign spouses are culturally distant but may offer other advantageous traits such as youth or better education. As such, legalization policies targeting the labor market do not only affect the behavior of immigrants but also the welfare of both immigrants and natives in the marriage market and subsequent generations through fertility choices.

Disentangling the role of legal status from other characteristics of spouses such as cultural differences is difficult, due to collinearity and the lack of exogenous variation. To overcome this identification challenge, we leverage 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. Therefore, the EU enlargements eliminated legal status benefits from the marriage market for new EU citizens living in other EU countries– but not for other immigrants. This episode constitutes an ideal research design for separately identifying the effect of legal status incentives and of cultural differences from other socio-economic characteristics. Focusing on Italy, we exploit registry data on the universe of marriages and separations in that country over the period 1998-2012. These data allow us to look at individual decisions involving 3.6 million marriages and over 200 thousand separations before and after the enlargements. We also supplement the analysis with data on couples' fertility to analyze changes in marital selection during the enlargement process.

The analysis proceeds in two steps, using first reduced-form techniques and then devel-

[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).

oping an equilibrium model of the marriage market, where marital, fertility, and separation choices are considered jointly. We first show, using difference-in-difference and event study techniques, that the enlargements of the EU had a profound effect on mixed marriages and separations as soon as the relevant information became known. The probability of intermarrying with native males decreases by 40 percent, ceteris paribus, for migrants from new EU countries relative to migrants from other countries after (the announcement of) the EU enlargements. The marginal increase in cohabitations does not offset this effect. Interestingly, 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 also entered the EU. During the same period, the hazard rate of separation for mixed couples formed before the enlargements between natives and citizens from new EU countries increased by 20 percent. Results on fertility are consistent with legal status having a primary role for both the quantity and the quality of marriages between immigrants and natives.

We then quantify the role of legal status through the analysis of a model that explicitly takes into account the trade-offs between legal status and cultural distance, as well as the general equilibrium effects that characterize marriage markets. Our model also allows us to study the welfare effects of reforms that grant legal access to migrants and to analyze the implications of highly debated and controversial immigration policies via counterfactual simulations.

We propose a transferable utility matching model, where individuals match on both observables and unobservables. In that process, individuals consider multiple observable traits such as age, education, and cultural distance but also whether the marriage confers legal rights to work to both partners. Trade-offs between these characteristics drive the marriage allocation in equilibrium. By altering these trade-offs, the enlargements therefore lead to a reallocation of spouses. Importantly, we consider a marital surplus that is specific to a given couple, by introducing a match-specific random component. Our specification allows for sorting on unobserved characteristics on both sides of the market, but also for complementarity between observable and unobservable characteristics, which allows us to match the substitution patterns induced by the policy we examine. We model fertility choices as a function of the marital surplus, to capture how changes in marital sorting induced by policy have demographic implications. The evolution of the random unobserved couple-specific component explains subsequent separation choices.

Our identification strategy exploits variation over time (before and after the EU enlarge-

ments) and across geographical areas with different population characteristics.[2] Specifically, we separately identify the contribution of different factors which simultaneously affect the marital equilibrium and later separation choices, by exploiting both observed matching patterns and separation rates over time and across markets. 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 value legal status. If new EU citizens are less keen to marry natives, the latter must in turn opt for other types of spouses, perhaps more culturally distant or with different characteristics. These trade-offs then reveals the value of cultural affinity and of different characteristics such as age and education.

Considering separations in addition to marriages allows us to disentangle the variation over time in marital gains parameters from potential variation in unobserved composition effects due to the EU enlargements. The intuition is that, while the marriage equilibrium is affected at the same time by changes in legal status acquisition and market competition in each period, separations are only affected by changes in legal status, keeping constant the observable traits of the spouse that led to marriage.

Our analysis provides evidence on the value of legal status, the absence of which leads to a penalty in the marriage surplus of about 20 percent, which reduces to 17 percent in areas with a large informal labor market. We also evaluate the role of cultural distance across different nationality groups, as revealed by marriage choices. We show that mixed marriages carry a marital surplus penalty that can be substantial across some nationality groups. We quantify various trade-offs along observables that show that the legal status effect on marital surplus is large enough to deviate the marriage allocation from positive assortative mating, explaining intermarriage choices. It is of similar magnitude to a 7-year difference between the age of spouses or about 6 years of education differences between spouses. We show that a one standard deviation change in the surplus leads to a change in fertility of about 3 percent. Our results suggests that the enlargements had a small effect on fertility overall, but changed the type of families in which children are born. We then quantify the welfare implications of the EU enlargements on marital surplus, and we show that it constitutes a transfer of welfare away from natives to migrants with different redistributional effects along gender lines.

We leverage our model to evaluate several counterfactual policies. The first one grants legal status to all immigrants in the country. Such generalized amnesties have been enacted in countries such as France, Spain, and partially in the United States (e.g. the 1986 Immigration

[2]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. Other examples of multi-market applications are Mourifié (2019) and Vickery and Anderberg (2021).

Reform and Control Act in the US). Although these policies aim to foster integration through access to the legal labor market, we show that they reduce mixed marriages in favor of homogamous ones by eliminating the benefits of legal status from the marriage market. Paradoxically, such measures may slow down the cultural integration of immigrant minorities, with potentially detrimental and persistent effects on successive generations (Bisin and Tura, 2019). Second, we evaluate the effects of an unconstrained open-border immigration policy by simulating a surge in migration inflows. Because of a marked asymmetry in cultural preferences along gender lines – intermarriages are more prevalent between native men and foreign women than between native women and foreign men – new arriving migrant women would get married while few migrant men do. Therefore, a large group of single foreign men will face difficulties in the labor market due to a lack of legal status and, more generally, in integrating into the socio-cultural host environment.

Our paper builds on and extends the previous literature on migration, marriage, and the role of culture. The first contribution is to show the implications of labor market policies directed at immigrants, such as legalization policies, on marital assimilation. Previous work has examined the effect of legal status on labor market opportunities (Amuedo-Dorantes et al., 2007; Lozano and Sorensen, 2011), 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.[3] Our results show how institutional changes targeted to the labor market affect society, both natives and immigrants, in a broader way through the types of families that are formed.

The second contribution of the paper is to emphasize the importance of cultural traits for marriage choices. Perhaps surprisingly, the cultural dimension in marriage has been overlooked,[4] despite its relevance for various social and economic outcomes 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; Fernandez, 2011; Alesina et al., 2013). Our work shows how cultural traits and economic policies interact in societies with increasing

[3]One notable exception is Azzolini and Guetto (2017), who document the negative effect of the EU enlargements 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 enlargements on the number of marriages but also on their characteristics and stability and on cross-nationality substitution in marriage markets. In addition, we also estimate the impact of legal status acquisition on separations and fertility.

[4]The previous literature on marriage has focused on the sorting mechanisms along age (Choo and Siow, 2006b; Choo, 2015; Shephard, 2019); income and body mass index (Chiappori et al., 2012); educational attainment (Chiappori et al., 2009, 2017; Gayle and Shephard, 2019); human capital and fertility (Low, 2019); and personality traits (Dupuy and Galichon, 2014). Exceptions are Ahn (2018) who studies cross-border marriages in East Asia, and Bisin and Tura (2019) who study intergenerational patterns of assimilation.

cultural diversity, with enduring effects, as marriages between immigrants and natives are a marker of cultural assimilation (Gordon, 1964; Kalmijn, 1994, 1998; De Graaf and Kalmijn, 2001; Algan et al., 2012) and can also accelerate this process (Meng and Gregory, 2005; Furtado and Trejo, 2013).

Finally, we add to the literature on marriage. From a methodological point of view, the paper advances the literature on marital matching, pioneered by Choo and Siow (2006b) who first provided identification of matching patterns under the assumption of separability of the unobserved heterogeneity component. This allows for closed-form tractability that generated a large literature with further methodological developments (Dupuy and Galichon, 2014; Galichon and Salanié, 2021; Ciscato et al., 2020). However, separability imposes strong restrictions on substitution patterns across different types of spouses (Decker et al., 2013).[5] By relaxing the additive separability, our model is more flexible and does not impose symmetry restrictions on substitution patterns. Independently, Chiappori et al. (2019) present a marriage model along educational lines, and, similarly to us, they allow for a couple-specific preference shock. Solving the model via simulations, they compare it to the separable Choo and Siow benchmark model. We also enlarge this literature by estimating jointly marriage and intra-household decisions regarding fertility and separations, considering multidimensional traits. The estimated model provides a rich characterization of the social and demographic consequences of immigration policies.

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 from marriage and on separations. Section 4 develops the multidimensional equilibrium model of the marriage market and subsequent separations, its identification and our estimation strategy. We present the estimation results in Section 5, as well as welfare and counterfactual analysis. Finally, Section 6 concludes.

[5] 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 lead to the same total surplus. This allows us to translate the individual matching problem into a matching along observables only. With a different approach, Galichon and Salanié (2017) and Galichon and Salanié (2021) 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. Recently, Gualdani and Sinha (2022) estimate a nonparametric model and discuss the consequences of the logit assumption imposed by the Choo and Siow model.

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. Figure 2 shows the composition of the 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 admission of immigrants from all other countries is instead regulated by a rigid quota system. The main pathway to an official stay is through 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, and 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 the 12 EU countries in the Strictness of Immigration Policy index produced by the Fondazione Rodolfo Debenedetti (Fumagalli and Boeri, 2009).

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. 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 3 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. 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 EU enlargements

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. Starting in the following decade, 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 ($EU2004$ 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 ($EU2007$ henceforth).

Both rounds of enlargement were preceded by complex and lengthy negotiations. Figure 4 shows the timing of the official approval and implementation. The EU Council – the executive branch of the Union – agreed upon the 2004 enlargement in December 2002; the timing of the 2007 enlargement is similar, the EU Council taking the final decision in December 2005.[6] In practice, the timing and extent of legal status concession were surrounded by a high degree of uncertainty, as the single member countries retained significant discretion in the implementation of EU directives. For instance, France, Germany and the United Kingdom all maintained significant restrictions to free mobility and the labor market access of new EU citizens in the first years after enlargement. Italy allowed free mobility and full labor market access, but such decision came after a heated debate, and the government decrees implementing the EU directives were approved only a few days before the official date of the enlargements (e.g., December 28, 2006, for the enlargement of January 1, 2007). Therefore, new EU citizens faced large uncertainty regarding the residence and working rights in Italy – like in most other countries – until the very last days before the enlargement. For this reason, we will allow for anticipation effects between the approval and implementation dates in our empirical analysis.[7]

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

[6] Appendix B provides additional details on the enlargement process.

[7] Media coverage of EU enlargements in Italian newspapers confirms that uncertainty about the actual extent and features of legalization unraveled only in the last days before the official enlargements dates; see Appendix C.

transitory regimes imposing barriers against new EU citizens from Eastern Europe. For instance, Germany eliminated entry quotas for Polish workers only in 2011. 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 the 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 on dependent employment in most economic sectors.

2.3 The effect of the EU enlargements on labor market outcomes

Free access to residency and work improved the labor market opportunities of new EU citizens. To gauge the magnitude of such effects, we would ideally compare employment and income before and after legal status acquisition. Such comparison is complicated by the fact that information on irregular immigrants is typically not available from official sources. However, since 2001 the Institute for Multi-ethnic Studies (ISMU) has been conducting an annual survey on a representative sample of about 8,000 migrants residing in the Italian region of Lombardy, including the irregulars. The sampling scheme and the questionnaire are specifically designed to elicit truthful reporting of legal status along with other individual characteristics, including employment status and income. For this reason, the ISMU survey has been extensively used in previous studies of irregular migration (see, e.g. Dustmann et al., 2017; Guriev et al., 2016).[8]

Table 1 shows that, on average, regular immigrants in the ISMU survey earn a substantially higher average (monthly) income than irregular immigrants – €798 and €562, respectively. Interestingly, the difference is entirely due to males, for whom legal status is associated with higher employment rates and to almost a doubling in income; for females, the wage premium to legal status is instead negative, though small in magnitude.

Differences in employment and earnings between regular and irregular (male) immigrants partly reflect differences along other individual characteristics, reported in Panel B of Table 1. In particular, regular immigrants are on average older, more educated, and more likely to be in a stable marital relationship. Comparing labor market outcomes over time between

[8]In Appendix D, we provide additional details on the survey and explain how we combine it with data on generalized amnesties of irregular immigrants to estimate the size of irregular migration in Italy.

new EU citizens and other immigrants, before and after the EU enlargements, allows us to identify the (intention-to-treat) effect of legalization on labor market outcomes separately from selection into legal status on other individual characteristics.

Figure 5 compares $EU2004$ and $EU2007$ immigrants, before and after they entered the EU, with all immigrants from other countries. This event study approach addresses the challenges to the estimation of difference-in-differences models with staggered treatment adoption, discussed e.g. by De Chaisemartin and d'Haultfoeuille (2020) and Goodman-Bacon (2021). Admission to the EU brings a reallocation of workers from the informal to the formal sector, leading in turn to an increase in earnings. Such effects are stronger and statistically significant for males, while the estimated coefficients have the same sign but are smaller in magnitude and not significantly different from zero for females.[9]

We next investigate the implications of the improvement in the labor market opportunities of (male) immigrants from new EU countries for marriages and separations with natives.

3 The effect of EU enlargements on marriages and separations

Figure 6 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). Panel (a) 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 thereafter. In contrast, marriages among natives exhibit a steady and slow decline over the same period. Panel (b) of the same figure 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 a divorce). Separations between new EU immigrants and natives exhibit abnormal spikes after the EU enlargements, and decline immediately after.

This preliminary evidence is consistent with legal status acquisition reducing benefits from marriage between migrants and natives. We next illustrate our strategies for precisely quantifying these effects.

[9] Appendix Table A2 reports the estimated magnitude of such effects over the six years post-enlargement. The table also shows that estimates remain virtually identical when controlling extensively for other individual characteristics, to account for changes in the composition of the immigrant population after the enlargements.

3.1 Marriages: data and methodology

The marriage rates depicted in Panel (a) of Figure 6 do not take into account changes in the number of available matching opportunities in the market. In particular, the dynamics of marriages between natives and new EU citizens reflects both the marital surplus of those marriages and the number of immigrants from new EU countries that are present in Italy at each point in time.[10] To take into account the change in the pool of available men and women, we follow the approach of Choo and Siow (2006b), who compute the *gains from marriage* between males from country x and females from country y as:

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

Intuitively, a higher number of observed marriages between type-x males and type-y females (the numerator of equation 1) relative to the number of potential matches between the same types (the denominator of equation 1) reveals more gains from marriage.

The gains from 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 the country of origin-province-year cells exploiting rich registry data made available by ISTAT through its Laboratory for Elementary Data Analysis (ADELE).[11] Vital statistics registries provide detailed information on the universe of marriages celebrated each year in Italy between 1998 and 2012 among adult individuals – more than 3.6 million marriages (see Panel A of Table 2).[12] Information on spouses' country of origin, reported in the marriage records, allows us to compute the numerator of the marital matching surplus $\Phi(x,y)$ in equation (1). Following Choo and Siow (2006a), we alternatively measure the numerator of $\Phi(x,y)$ by the number of cohabiting couples not in a legal relationship, or by the sum of marriages and cohabitations, as available from the decennial population Census.

Measuring the number of singles at the denominator of the marital matching surplus $\Phi(x,y)$ in equation (1) confronts two main challenges. First, the exact number of residents

[10] Previewing some of the results reported below, the evidence in Panel (a) of Figure 6 goes against sham marriages. First, the peak in marriages happens two years prior to the enlargements. Second, the increase observed before the EU enlargements is a result of the large inflows of immigrants into Italy. Furthermore, analyzing separation patterns leads to the same conclusions.

[11] 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.

[12] Since we focus on the Italian marriage market, our data include only marriages celebrated in Italy. We ignore immigrant couples married before arriving in Italy, as they are not part of the Italian marriage market.

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 2 show the number of observations and summary statistics for years 2001 and 2011, respectively. We then interpolate these data outside Census years exploiting variability over time of official migration inflows. This allows us to account for non-linear changes in the presence of singles between any two Census years. The exact procedure is described in Appendix D.

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 enlargements, as we discuss in detail in Section 3.2 and in Appendix D. 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 temporary residence and work permits to 700 thousand irregular immigrants – a third of the total immigrant population at that time.[13] Using administrative data on the universe of applications for amnesty, 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 amnesty applications across the same cells. Appendix D provides details on 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 from marriage $\Phi_{IT,y,t,p}$ between native males ($x = IT$) and females from country of origin y, in province p, and year t:

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

where $newEU_y$ is a dummy for females from $EU2004$ or $EU2007$ countries; $postEU_{y,t}$ is

[13] 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 status – as it was the case, instead, for new EU citizens. Devillanova et al. (2018) provide additional details.

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; FEs are different sets of fixed effects for year, province, and foreign spouse's country of origin, possibly interacted with time trends; finally, $\varepsilon_{IT,y,p,t}$ is a residual term summarizing the effect of other determinants of gains from marriage. We estimate an analogous equation for the gains from marriage between native females and males from country of origin x, $\Phi_{x,IT,p,t}$. Implicitly, by following Choo and Siow (2006b), we consider a static model. We discuss dynamic considerations referring to anticipation effects below.

The coefficient β in equation (2) captures the relative change in gains from intermarrying with natives after the EU enlargements between immigrants from new EU countries and other countries, respectively. Given that immigrants may acquire legal status 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 is an "intention-to-treat" effect).

3.2 The effect of the EU enlargements on marriage formation

3.2.1 Main results

Figure 7 plots the gains from marriage with natives, as measured by $\Phi(x, y)$ in equation (1), for different groups of immigrants during the period 1998-2012.[14] The announcements of both the 2004 and 2007 enlargements bring a decline in gains from marriage between new EU citizens and natives, with the exception of marriages between $EU2004$ males and native females. However, the latter includes less than 100 intermarriages per year, so gains from such marriages may be imprecisely estimated.

We quantify the changes in Figures 7 by estimating the difference-in-differences specification in equation (2); results weighted by province population are presented in Table 3.[15] Columns (1)-(3) show the results for marriages in which the husband is native, while columns (4)-(6) show the results for marriages in which the wife is native. According to the baseline specification (columns 1 and 4), which includes year, province, and country of origin fixed effects, admission to the EU brings a 0.8 log point decrease in gains from marriage between females from new EU countries and native males, and an even larger decrease – twice as much – in gains from marriage between males from new EU countries and native females. However, the two estimates become very similar, and close to unity, when we control for

[14] Figure A4 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 A1 reports the complete list of countries in each group.

[15] The non-weighted results are virtually identical and are presented in Appendix Table A3.

underlying trends by province and country of origin as well as for province × country of origin fixed effects (columns 2 and 5). According to these estimates, admission to the EU decreased gains from marriages with native males and females by 58 and 64 percent, respectively. The estimated effect is slightly larger when including only official immigrants in the computation of the number of singles (columns 3 and 6), consistent with the discussion in Section 3 and Appendix D. 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. Indeed, we also assess the robustness of our estimates to allowing for a progressively larger number of irregular immigrants, which has, however, little impact on our results.[16]

Figure 8 plots dynamic, year-specific effects of enlargements on gains from marriage between natives and, respectively, immigrants from $EU2004$ and $EU2007$ countries; see also the Appendix Table A4. 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 $EU2007$ females. This is consistent with native males substituting $EU2007$ for $EU2004$ females when the announcement of the 2004 enlargement lowers expected gains from 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 8 documents the absence of differential trends before the announcement of either enlargement.

3.2.2 Robustness

In Table 4 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 from marriage, percent changes Δ_x and Δ_y in the number of male and female singles, respectively, would induce a percent change $(\Delta_x + \Delta_y)/2$ in the number of marriages. This restriction results from functional form assumptions in Choo and Siow (2006b), by which the log number of squared marriages is rescaled by the number of singles. In columns (1) and (5) of Table 4, we allow for a more flexible specification by relaxing

[16] If we increase the number of irregular immigrants by as much as 300%, the estimated effect of the EU enlargements on gains from marriages between native males and foreign females changes only from -0.86 to -0.64, and by even less for couples between native females and foreign males; in both cases, the estimated effect would remain strongly statistically significant. These results are presented in Appendix Figure A1.

linearity in the scaling factor involving the singles, similarly to Mourifié (2019); Mourifié and Siow (2021). Second, the propensity of foreigners to intermarry with natives likely reflects the availability of other singles from the same country of origin. Indeed, the period after the EU enlargements witnessed large inflows from new EU countries, so the estimated decrease in the number of intermarriages could reflect a greater availability of singles from the same country of origin. In columns (2) and (6) of Table 4, we thus control for the log number of singles from the same country of origin of the foreign spouse, in each province and year. In columns (3) and (7) we include all these additional 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 virtually identical to the baseline estimate in Table 3. In the most stringent specification (columns 4 and 8), the effect of the EU enlargements on gains from marriage amounts to a full log point. Put differently, the EU enlargements decreased the number of marriages between natives and new EU citizens by approximately 40 percent (keeping constant the number of potential matches in each marriage market, as measured by the distribution of singles across provinces).[17]

An additional concern is that the timing of approval and announcement of the 2004 EU enlargement coincides with the implementation of the 2002 generalized amnesty of all irregular immigrants in Italy (2002-2003). The latter policy should have also reduced gains from marriage between natives and all immigrants in Italy. Figure 7 shows, indeed, that gains from marriage with natives decline for all nationalities between 2002 and 2003. However, the decline observed for other nationalities is generally smaller and more temporary than the one observed for $EU2004$ citizens, consistent with the fact that (differently from the EU enlargements) the amnesty only granted temporary permits. Most importantly, since immigrants from other nationalities constitute the control group in our main analysis of new EU citizens, any effect of the amnesty should bias downward the magnitude of our coefficient of interest. In practice, allowing for an effect of the 2002 amnesty on all immigrants has no discernible effect on our estimates; results available upon request.

Results are also robust to re-estimating the effect assuming that the whole of Italy is a single marriage market; to different marriage market partitions;[18] to leaving out one country

[17]The number of singles is by construction endogenous (Mourifié, 2019). In Appendix Table A5, we instrument the number of single males and females by lagged immigrant population vectors (supplies) of males and females (Mourifié and Siow, 2021). The coefficients of interest remains identical to our baseline estimates.

[18] In our analysis we assume that markets are separate and local, i.e., individuals match within the province's marriage market where they live in. In the data, spouses share the same province of residence in more that 83 percent of our marriages. Nevertheless, we assess the sensitivity of results to allowing for a different partition of marriage markets, namely by Local Labor Markets (LLM). Results are very similar when using one or the other definition; see Appendix Table A6.

of origin at a time; to account for sample selection potentially induced by the fact that there are no marriages in some cells (so the logarithm in equation (1) is not defined); and to alternative one-way and two-way clustering of standard errors. These results are available upon request. Finally, in Table A7 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 effect of interest does not decline with greater diffusion of broadband Internet.

3.2.3 Effect on Cohabitation

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, we replicate our analysis by exploiting information on cohabiting unmarried couples, as available from the population census in 2001 and 2011.[19] If our effect of interest was entirely (or partly) driven by substitution of official marriages with cohabitations, we should observe a positive effect on cohabitations alone, as well as a zero (or weaker) effect on the total number of couples living together (i.e., the sum of married and unmarried couples).

However, the results in Table 5 show that the enlargements actually affected the total number of couples formed between natives and new EU citizens. Even though official marriages with new EU citizens are partly replaced by cohabitations after the enlargements, this substitution is too limited to alter substantially our main conclusion. Specifically, columns (1) and (4) of Table 5 report a negative and significant estimated coefficient for official marriages formed in the restricted sample of years before the 2001 and 2011 censuses. By focusing on cohabitations alone, we observe that the EU enlargements lead to an increase in gains from cohabitation between native males and females from new EU countries (column 2), while the effect on cohabitations between native females and males from new EU countries is not statistically significant (columns 5). Most importantly, when adding cohabitations to official marriages in the numerator of the gains from marriage, our effect of interest is still negative and sizable in magnitude (columns 3 and 6).

[19] Cohabitations identify heterosexual love relations, as self-declared by the respondents, between partners who reside together. We do not have information regarding the starting date of the cohabitation. In order to retrieve a flow dimension, we thus select cohabitations of natives with migrants permanently residing in Italy in the last five years before the two censuses, i.e., from 1997 to 2001 for the 2001 census, and from 2007 to 2011 for the 2011 census. The sample is much smaller than in the main analysis because cohabitations data only covers two years rather than 15.

3.2.4 Matching on Other Characteristics

So far, we have documented how successive EU enlargements have affected the number of mixed marriages. In addition to cultural origins, individuals also match on other observable characteristics, such as age and education. Figure 9 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 EU enlargements). 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 wives by two to four years. The gap increases considerably for intermarriages between native husbands and foreign wives (second graph), while the opposite is true for intermarriages formed by native wives and foreign husbands (third graph). Turning to education, wives are slightly more educated than husbands (fourth graph). The gap widens for intermarriages between native husbands and foreign wives, while it is reduced for intermarriages between native wives and foreign husbands.

This shows that sorting in the marriage market is multidimensional and that 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. The Appendix Table A8 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 enlargements.

In Section 4, we incorporate these trade offs into a multidimensional matching model of the marriage market. Before doing that, we present the evidence on separations.

3.3 Separations: data and methodology

ADELE provides access to rich administrative data on separation records from civil courts' registries. Separation is the formal act that starts the divorce proceeding. Until 2015, a minimum period of 3 years of legal separation was required in order to eventually obtain a divorce. For this reason, separations provide a more accurate representation of the timing of marital instability.

Our data cover the universe of separations in Italy from 1998 to 2012, including detailed information on separation proceedings, date and place of marriage, and date and place of birth for both spouses for a total of over 200 thousand separations over the period 1998-2012 (see Panel B. of Table 2). We follow each married couple over time until potential separation. Following 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 be unrestricted, we can estimate γ by partial maximum likelihood. This approach remains very tractable while allowing for considerable flexibility in the baseline hazard rate $\lambda(d)$. We also 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, systematic changes (if any) in the risk of divorce for mixed couples after two years of marriage, i.e., when the foreign spouse can apply for Italian citizenship.[20] As for our main effect of interest, the effect of EU enlargements is identified off differential changes in the hazard rate of separation between new EU citizens and other immigrants, before and after the enlargements.

3.4 The effect of EU enlargements on separations

Table 6 shows the estimated effect of EU enlargements 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* and its interaction with dummies for different periods after the enlargements (in addition to calendar year fixed effects). We distinguish between the first year after the enlargements and the following years, respectively, because the decrease in gains from marriage should produce an immediate increase in separations after the enlargements. 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 enlargements. 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 enlargements are particularly stable. 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 specifica-

[20] We do not detect any systematic shift in the hazard rate of separation after two years of marriage, see Figure A2. This is likely due to the fact that, while foreign spouses can apply for citizenship after two years of marriage, there is considerable uncertainty regarding the duration of the procedure and the final outcome.

tion 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).

3.5 Fertility

We provide evidence on fertility patterns using matched registry data on marriages, separations, and the universe of birth records for the period 1998-2014; see Appendix E for additional details. We investigate the impact of legal status acquisition on fertility by estimating specification (2) after replacing, on the right-hand side, a dummy equal to 1 for couples having a child within 3 years from marriage. The results, reported in Table 7, show that fertility increases relatively more (by an additional 5-10 percentage points) for couples formed between new EU citizens and natives, after the EU enlargements, than for other heterogamous couples. To the extent that fertility proxies for higher marital surplus, as it is typically assumed (see, e.g., Becker, 1973, 1974), these findings are also consistent with legal status playing a primary role for both the quantity and the quality of marriages between immigrants and natives.

4 Multidimensional marital matching model

In this section, we develop a model of marital matching, fertility, and separations that allows for the trade-offs studied above and provides a coherent framework to quantify the effect of legal status and cultural differences. We next discuss the identification of the model and the econometric methods used for the estimation.

4.1 Setup

We consider a transferable utility (TU) marital matching model. 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. The marriage market is frictionless, i.e., any individual has complete and costless information about all subjects present in that market and their characteristics.

We consider 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.[21]

Each marriage market is a two-sided market, with a population of men denoted by $i \in I$ and women $j \in J$. Men and women are heterogeneous along a vector of relevant attributes for the marriage market. In particular, each man i is characterized by a vector of observable attributes $x_i \in X$, while each woman j is characterized by a vector of attributes $y_j \in Y$. In our empirical analysis x_i and y_j represent a bundle of personal and socio-economic characteristics, comprising respectively age, educational attainment, geographic area of origin, and wealth (proxied by home-ownership):[22]

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

Given I men and J women, $I \times J$ matches are potentially observed in each market p at 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, so the matching measure needs to satisfy 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, in period t and province p. Denote $\omega_{tp}(x_i)$ and $\omega_{tp}(y_j)$ the outside options of remaining single for a man or a women in period t and province p. We let the outside value depend on the

[21] We consider 21 different provinces (Ancona, Bari, Bologna, Brescia, Cagliari, Catanzaro, Firenze, Genova, L'Aquila, Milano, Napoli, Padova, Palermo, Perugia, Potenza, Roma, Salerno, Torino, Trento, Trieste, and Venezia). We assume that markets are separate and local, see footnote 18 for further details.

[22] Home-ownership is not recorded in marriage registries, but this information is recorded in the Census, together with the other characteristics in x_i or y_j. Census data, thus, 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 $\bar{y}_j = \{y_j^A, y_j^{Ed}, y_j^{Orig}\}$ are the set of characteristics excluding homeownership. In our estimation method, we use those probabilities to integrate out housing when computing marriage rates. This allows for another determinant for migrants to marry natives, which could confound the role of legal status.

geographic area of origin of the spouses, the time period and the geographic origin of the individual. 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 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 is the solution of the following welfare maximization problem over all potential matches in the market, including the 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), which is unique under mild continuity and compactness conditions of the surplus. A matching is stable if nobody would prefer to deviate from the assignment, i.e., there are no blocking pairs (Chiappori and Salanié, 2016). We solve the optimal matching from the maximization problem in (5) numerically, given the absence of a closed form solution to our model. We detail this below.

4.1.2 Surplus specification

We characterize the joint marital utility to include both systematic and idiosyncratic components. In what follows, we provide a flexible parametrization for the surplus function. We choose a log additive structure for its various components. This choice is made as we observe in our data that changes in the legal status affect couples with low but also high surplus as measured by age and education, a feature that an additive specification would struggle to reproduce. The surplus is represented as follows:

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

The first part of the joint marital surplus comprises four systematic components denoted by $\phi_t^k(x_i, y_j)$ 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, at time t. 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 single individuals own property or

not. We assign individual cultural identities based on $N_{Orig} = 8$ geographic areas, namely Italy, EU15, EU2004/07, Other Europe, Asia, Africa, South-America and Other OECD countries. We denote $I_{x^k=g,y^k=g'}$ an indicator variable equal to one if the k-th characteristic of both partners belongs to the group g and g' respectively. We parametrize the $\phi_t^k(x_i, y_j)$ components of the surplus as follows:

$$\phi_t^k(x_i, y_j) = \sum_{g,g'}^{N_k} \alpha_{t,g,g'}^k I_{x^k=g,y^k=g'} \ , \quad k = \{A, Ed, Orig, W\}, \tag{7}$$

where the parameters $\alpha_{t,g,g'}^k$ account for the strength of mutual attractiveness across spouses' observable characteristics. In particular, we interpret the parameters $\alpha_{t,g,g'}^{Orig}$ as a measure of the cultural affinity between individuals from countries g and g'. We allow the $\alpha_{t,g,g'}^k$ coefficients pertaining to geographical origin to change over time to account for potential variation in preferences due to differences in selection into migration following the EU enlargements.[23]

The next component of the surplus in (6) is $\beta_{tp}(x_i, y_j)$, which represents the value of legal status obtained through marriage. 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 marriages where both spouses have legal access to the labor market when single. This includes homogamous marriages between Italians or EU15 citizens before the enlargements, homogamous marriages of citizens from EU2004/07 after the enlargements, but also mixed marriages between Italians or EU15 citizens before the enlargements and with those from the new accession countries after the enlargements. For marriages where neither spouses have legal access to the labor market through their citizenship, we assume that the surplus is lower by a factor $\beta_{tp}(x_i, y_j) = \underline{\beta}_p \leq 1$, which is potentially province dependent. We assume that our parameters $\underline{\beta}_p$ vary across provinces as a linear function of the share of the shadow economy observed on informal labor market data.[24] It is possible that in the South of Italy, where informal work is more prevalent, legal access may not be worth as 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 does not (e.g. a mixed marriage between 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 on whether the spouse with legal status is male or female.

[23] Our parametrization is fully flexible, and it does not collapse the multidimensional space of characteristics into a single index aggregating the observables in such a way that all individuals share similar preferences about the other side of the market (as in Chiappori and Oreffice, 2008; Chiappori et al., 2012).

[24] The shadow economy is measured by the estimated fraction of irregular workers over total workers; see Istat Shadow Economy Dossier (2010), Italy.

Finally, the surplus depends on a match specific random component, $\varepsilon_t(i,j)$, which rationalizes 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 –"love"– on both sides of the market. We interpret it as a random love component, which is observed by the individuals, but unobservable from an econometric perspective.[25] We assume it follows an exponential normal distribution, with zero mean and variance σ.[26] In this respect, our model represents a considerable departure from the literature originating from Choo and Siow (2006b), which relies on individual and separable idiosyncratic shocks over observed characteristics of mates only. Such a separability assumption leads to a tractable model and offers a transparent identification of the gains from marriage. However, it also rules out complementarity between unobserved characteristics (Chiappori and Salanié, 2016), and imposes symmetry restrictions on the substitution patterns across different types of spouses (Decker et al., 2013). In this respect, our model is more flexible, as we show in the next Section.[27] An additional advantage of having a couple-specific unobserved component is that it allows for separations through the evolution of $\varepsilon_t(i,j)$, as detailed below.

4.1.3 Fertility

We link the marriage decision to subsequent fertility by assuming that it is determined by the marriage surplus and a random, normally distributed shock at the couple level. Denote $C_{tp}(i,j)$ an indicator function equal to one if the couple has at least one child within three years from the date of marriage. The probability of having children within that time window

[25]The idiosyncratic component might also be interpreted as a non-monetary return from marriage. A different approach to rationalize the heterogeneity in marital sorting is introducing frictions in the market and explicitly modeling a meeting technology. Frictions imply that any individual has imperfect and costly information about potential mates in the market. Randomness on the meeting technology guarantees that similar agents, in equilibrium, will match different types of partners (Chiappori and Salanié, 2016).

[26]While the variance of the shock is constant across matches, the variance of the surplus is match specific as the shock enters the surplus in multiplicative way, leading to heteroskedasticity. Heteroskedastic models have been estimated by Chiappori et al. (2017) and Galichon and Salanié (2021) for instance, with an improvement in the fit and a rejection of constant complementarities. Note that the assumption of a normal distribution for $\varepsilon_t(i,j)$ ensures continuity for the surplus function.

[27]Chiappori et al. (2019) present a unidimensional model, and, similarly to us, they allow for a couple-specific preference shock. Solving the model via simulations, they compare it to the separable Choo and Siow benchmark model. They show that the supermodular core is robust to the inclusion of interaction terms, point estimates of the joint surplus are biased upwards, and the bias from miss-specification grows slowly with the magnitude of the contribution of the interaction terms (see also Galichon and Salanié, 2021), in line with our discussion. More recently, Gualdani and Sinha (2022) estimate a nonparametric model and discuss the consequences of the logit assumption imposed by the Choo and Siow model.

is written as:

$$Prob\left(C_{tp}(i,j)=1\right)=\mathcal{F}\left(\gamma_0+\gamma_1\Phi_{tp}(i,j)\right) \tag{8}$$

where $\mathcal{F}$ is the cumulative normal density. We assume that single individuals do not become parents. Policy changes that impact the marriage market, as the access to legal work we are studying, may change fertility patterns through a change in the surplus of already-formed couples or by reallocating spouses. We explore these implications below.

4.1.4 Divorce

Although our marriage model is static, we introduce a possibility of divorce. Modeling divorce choices adds an additional outcome related to legal status acquisition and allows us to improve on our identification strategy.

Divorce results from changes in marital surplus over time, as driven by policy shocks, captured by $\beta_{tp}(x_i, y_j)$, or by changes in the couple-specific "love" shock $\varepsilon_t(i,j)$. Both shocks are unanticipated at the moment of marriage. However, we model the "love" shock as a random walk, so the best expectation of $\varepsilon_{t+1}(i,j)$ is $\varepsilon_t(i,j)$, observed at marriage:[28]

$$\varepsilon_{t+1}(i,j)=\varepsilon_t(i,j)+\eta_{t+1}(i,j), \qquad \eta_{t+1}(i,j)\sim\mathcal{N}(0,\sigma^2_\eta). \tag{9}$$

Negative surprises and changes in legal status potentially trigger divorce, in line with Becker-Coase theorem (Chiappori et al., 2015), but we assume that only a fraction $1-\delta$ of couples separate when the surplus becomes negative.[29]

4.2 Identification

The model we develop above has two sets of parameters, some that can be identified in a single cross-section of observed marriages and some that require the observation of at least two cross-sections with a relevant policy change. We discuss them in order.

The first set of parameters consists of the surplus linked to observable traits $\{\alpha^k_{tgg'}\}$, the outside value $\{\omega_{tp}\}$, and the standard deviation of the "love" shock, σ. The parameters $\{\alpha^k_{tgg'}\}$ drive the optimal allocation of spouses across groups defined by traits k. For instance, if α^k_{tgg} is larger than $\alpha^k_{tgg'}$ the resulting marriages would display homophily. Hence the degree of assortativeness in a given market would pin down this set of parameters. For identification purposes, because the primal problem in (5) is scale-invariant, we introduce

[28] Voena (2015) assumes a similar stochastic structure when modeling divorce.

[29] We assume that δ is constant over time given the short time span we are looking at.

the following coefficient restrictions. For homogamous couples, we impose $\alpha^k_{t,1,1} = 1$ for all $k = \{A, Ed, Orig, W\}$ at $t = t_1$. The other coefficients $\alpha^k_{t,g,g'}$ are left free to vary to match the marriage rates for the other groups. Thus, the surplus for a couple of Italians, for example, with high education, both aged 22 and with no housing, residing in province p at time t is equal to $1 - \omega_{tp}(x_i) - \omega_{tp}(y_j)$. Given the parametrization above, then the parameters ω_{tp} are pinned down by observing the marriage rates of such spouses in province p at time t.

The outside value changes the fraction of individuals who marry, but not the optimal allocation of spouses. In contrast, an increase in the variance of the "love" shock will lead to an increase in the total number of marriages, but also to a reallocation of spouses. The latter property comes from the fact that the "love" shock enters the surplus in a multiplicative way and those with a higher $\alpha^k_{tgg'}$ benefit more from a larger "love" shock. Hence the variance of the "love" shock and the outside value have different implications on the marital outcomes and can be recovered separately. Moreover, we restrict the variance to be the same across periods and provinces, while we let the outside option vary to achieve further identification. Appendix F discusses the identification of a simplified version of the model with only one cross-section of marriage data and we also show through a Monte Carlo experiment, that each of these parameters is identified in cross-sectional data across different markets.

The second set of parameters is related to legal status $\beta_{tp}(x_i, y_j)$. They are not identified separately from the $\{\alpha^k_{tgg'}\}$ parameters in a single cross-section due to collinearity. However, a change in the legal status of one group over time, in the form of the policy that we evaluate in this paper, is enough to identify the two sets of parameters separately by contrasting two periods before and after the enlargements. Figure 10 provides an intuition for a simplified framework. For instance, observing marriage patterns after the enlargements allows us to identify cultural affinity coefficients alone for marriages formed by natives and citizens from the enlargement countries. Turning to the before period, we identify the product of cultural affinity and the legal status penalty for that group. Contrasting the two periods allows us to identify the penalty for legal status. The longitudinal variation, in turn, allows us to recover the surplus associated with a marriage between Italian men and women from other countries, even though that group is not directly affected by the policy that changes legal status.

The difference-in-differences design, just described, identifies model parameters under the assumption that the surplus related to cultural affinity is constant over time. This assumption might be too strong in our case as the enlargements could have attracted migrants with different unobservables or that are perceived differently by natives. To account for such changes in cultural affinity, we also rely on separation data as explained below. We focus on three cohorts, $c \in \{c_1, c_2, c_3\}$, according to the timing of marriage, 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, hence before the enlargements. The second cohort, c_2, consists of couples formed between Italians and EU2004 in 2002-03 and potentially separating by 2004-05 and marriages formed between Italians and EU2007 in 2005-06 and potentially separating by 2007-08. In this second cohort, citizens from the accession countries may reconsider their initial marriage choice once they have legal access to the labor market granted through the enlargements of the EU. The third cohort, c_3, consists of marriages formed between Italians and EU2004 in 2005-06 and potentially separating by 2007-08 and marriages formed between Italians and EU2007 in 2008-09 and potentially separating by 2010-11. Cohorts c_1 and c_2 both have a surplus related to the cultural affinity that is formed in periods before the enlargements, but c_2 in addition experience the change in legal status. Differences in separation rates across cohorts allow us to isolate the causal effect of legal status. Cohorts c_1 and c_3 do not experience the change in legal status, but the surpluses related to cultural affinity are formed in different periods, and their comparison identifies potential differences in selection between the two periods.

Finally, we identify the parameters related to fertility in equation (8) by exploiting the observed variation in fertility patterns before and after the policy change across cultural groups, conditional on observables, as in the difference-in-differences estimation of Section 3.5.

4.3 Estimation

Once restrictions are imposed, we have a total of 118 parameters, 6 of which pertain to legal status and 30 to the contribution of cultural affinity. The remaining parameters characterize the value of being single (30), the contribution of age to the marital surplus (35), the value of housing (6), education (3), fertility (4), and separations (4).[30] Let $\boldsymbol{\theta}$ denote the vector of parameters. We estimate the model parameters via a Method of Simulated Moments (MSM) estimator, by matching the vector of moments observed in the data, $\widehat{\Lambda}$, with their theoretical counterparts predicted by the model via simulations, $\Lambda(\boldsymbol{\theta})$.[31] We use a Nelder-Mead algorithm with various initial values to check for robustness. Given a weighting matrix

[30] The outside value is parameterized in a parsimonious way: $\omega_{tp}(x_i) = \omega_p^1(1 + I_{t=t_2}\omega^2) + \sum_j \omega_j^O I_{Orig_i=j} + \omega^S * ShareSingle(p, Orig_i)$, where $ShareSingle(p, Orig_i)$ is the province and origin specific share of either single men or women. The specification imposes a common trend to all provinces, parameterized by ω^2. Note that, in the model, marital outcome trends in each province are determined by the outside value, but also by the distribution of characteristics of singles that are province and time specific, so that the resulting marital outcomes may not display similar trends across provinces.

[31] For a more detailed explanation of the estimation method through simulations and its implementation, we refer the reader to Adda and Cooper (2003).

$\boldsymbol{W}$, the estimated surplus parameters are obtained as:

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

From data, we recover matching patterns over observable characteristics $\bar{x}$ and $\bar{y}$, $\hat{\mu}_{tp}(\bar{x}, \bar{y})$, by summing up all men i of characteristics $\bar{x}_i = \bar{x}$ matched to women j of characteristic $\bar{y}_j = \bar{y}$, in province p at time t.[32] Those observed marriage rates are used as observed moments in the estimation. We add four moments related to fertility. We first use the difference-in-differences estimates in Table 7, columns 4 and 8, i.e., how the reform affected the fertility of couples formed by natives and foreigners, and specifically citizens from the enlargement countries. We add to these two moments the average probability of having children for these two groups so that the model also matches levels. We collect those moments in a vector $\hat{\gamma}$.

We also use information on separations, but given that they are infrequent in Italy, we do not use separation rates for all possible types of marriages. Instead, we consider more aggregate moments. Those data moments characterize divorce patterns of three different cohorts of married couples that we follow through time as explained in the previous section. For each of these three 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 ζ_c on the observed data from:

$$D_{cij} = Z_{cij}\zeta_c + u_{cij}. \tag{10}$$

The vector of observed moments consists of 387,072 marriage rates across all characteristics, provinces and time periods, 21 moments describing the probability of separation and four moments describing fertility choices:

$$\widehat{\Lambda} = \{\hat{\mu}_{tp}(\bar{x}, \bar{y}), \hat{\zeta}_c \hat{\gamma}\}, \quad t \in \{t_1, t_2\}, \;\; c \in \{c_1, c_2, c_3\}, \;\; p = 1, \ldots, P.$$

Theoretical moments are computed as follows. We recover marital matching moments through simulations. For each period t and province p, we simulate N marriage sub-markets

[32] We also account for the potential omission of homogamous foreign marriages celebrated abroad and, thus, not recorded in marital registries. See Appendix G for more details.

of I men and J women. The distribution of individual characteristics in the simulated sub-markets 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.[33] 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 the mathematical programming solver Gurobi.

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. The second pool is 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 us to have a boost sample of foreigners, and to reduce the number of simulations. We derive marital matching patterns, $\mu^{\boldsymbol{\theta}}_{tp}(\bar{x}, \bar{y})$, by recording as married any allocated match with a positive surplus and consequently derive the theoretical distribution of single individuals, $\mu^{\boldsymbol{\theta}}_{tp}(\bar{x}, 0)$ and $\mu^{\boldsymbol{\theta}}_{tp}(0, \bar{y})$.

We consider $N = 6$ sub-markets per province p, with size $I = J = 600$, half of them with foreigners. This amounts to considering 75,600 simulated men and women per period, leading to over 45 million potential matches. Increasing N leads to more stable results as the sample size increases and more marriages between rarer types are observed. A larger N increases the computational time linearly. The choice of I is undoubtedly arbitrary but fixed both by considering computing time and the stability of the estimates. Increasing I increases the computing time in a quadratic way. It also implies that each individual i faces more potential spouses and, therefore, more shocks $\varepsilon_t(i, j)$, hence having a larger chance of finding a suitable match. However, once the value of remaining single is free to adjust with I, we observe that for a large enough I there are few changes in the overall allocation. Our choice of I reflects this regularity and allows for a tractable computing time.

From those marriages, we compute fertility as implied by equation (8). We construct the model counterpart of the separation rates defined in equation (10) by simulating three different cohorts followed before, during or after the enlargements, as explained in Section 4.2. For a given choice of $\boldsymbol{\theta}$, we estimate the auxiliary parameters $\zeta^{\boldsymbol{\theta}}_c$ from model (10) on simulated data, separately for each cohort c.

[33] We recover population vectors, $m_{tp}(\bar{x})$ and $f_{tp}(\bar{y})$, in order to satisfy standard accounting constraints. A graphical representation of the distribution of observed traits by gender is reported in Figure A3.

5 Results

This section first discusses the model's fit and parameter estimates. We then present the estimated trade-offs between different dimensions; passing then onto a discussion of the welfare effects of the policy, and finally to various counterfactual simulations.

5.1 Model fit and results

Figure 11 displays the fit regarding marriage patterns. The correlation between the observed and predicted marriage rates is equal to 0.84 before and after the enlargements.

The model's predictions regarding fertility and separation rates for cohorts followed before or during the enlargement process are displayed in Table 8. The model captures the general patterns of fertility both in levels and in the marginal change, and separation across cohorts and countries of origin, even though few parameters that we estimate directly impact these outcomes in the model.

The coefficients regarding legal access are displayed in panel A of Table 9. For a migrant, the penalty for having a spouse from the same origin rather than a native is the ratio $\bar{\beta}_p^{\male}/\underline{\beta}$ or $\bar{\beta}_p^{\female}/\underline{\beta}$ for foreign men or women, in turn. This ratio represents a reduction in the marital surplus of between 17 and 20 percent and is larger for foreign women. This penalty also displays geographical heterogeneity with a smaller effect in areas with a larger shadow economy, as informal work makes legal status less valuable. The gender differences in the penalty reflect the lower rate of mixed marriages where the wife is native, documented in Figure 3, and examined in Section 3.

Panel B of Table 9 reports the average outside values by geographical areas and time periods. Differences in those values are small across provinces, but the outside value increased over time, as fewer marriages were celebrated.

Panel C of Table 9 displays the effect of cultural affinity on the marriage surplus. This effect derives essentially from revealed preferences by observing the choice of spouses across periods and markets. The coefficient for homogamous marriages (along the diagonal) is assumed to be one. We find that cultural affinity among mixed couples is lower than homogamous ones.[34] The penalty for cultural heterogamy ranges from one percent (for couples involving Italians and other EU15 citizens) to 44 percent for couples formed of citizens of Asian and African countries. We find evidence of an asymmetric surplus as couples formed of Italian males with foreign women have a higher surplus than couples with Italian females

[34]Similar evidence arises in the speed dating experiments (Fisman et al., 2006, 2008).

and foreign males.[35] We refer the reader to appendix G.2, which shows that our cultural affinity parameters correlate with commonly used cultural distance measures such as genetic or language distance.

We find strong preferences for positive assortative mating along age, education, and wealth dimensions. Our estimates mirror the evidence presented in the literature (among many others, Choo and Siow, 2006b; Hitsch et al., 2010; Siow, 2015; Chiappori et al., 2012; Dupuy and Galichon, 2014; Chiappori et al., 2017; Galichon and Salanié, 2017; Low, 2019). We refer to Appendix G.1 for a description of the parameters relating to age, education, and wealth differences.

We further investigate the substitution patterns implied by our model. Denote $R_{xx'}$ the percentage marginal effect of the number of males of type x in the market on the number of single males of type x'.[36] Figure 12 plots the elasticity $R_{xx'}$ as a function of $R_{x'x}$. Decker et al. (2013) show that, in the Choo and Siow (2006b) model, the functional form imposes symmetry: $R_{xx'} = R_{x'x}$, represented by the red line. Instead, the elasticities derived from our model do not align on the 45-degree line. This degree of flexibility is especially relevant in our context characterized by a large majority and small heterogeneous minorities, so the effects of natives on minorities are hardly equal to those of minorities on natives. Some of the counterfactuals we present below explore changes in the size of minorities.

The estimated parameters regarding fertility, reported in Appendix Table G2, imply a small marginal effect of marital surplus on the probability of having children. A one standard deviation change in the surplus leads to a change in fertility of about 3.6 percent. Therefore, our results show that the enlargements had a small effect on fertility overall. However, they engendered sizable redistributive effects that we explore below.

5.2 Trade-offs: Legal Status and other Characteristics

The estimated model allows us to quantify the trade-off between different traits in response to the change in legal status. We take the perspective of a woman from the EU2004-07 countries and evaluate the surplus when she marries men differing by nationality, age, or education. Figure 13 plots the results. Panel A starts by looking at the surplus of women from the EU2004-07 countries marrying men of different origins before and after the EU enlargements. It displays the surplus formed by combining the components related to

[35] In addition, we find a minimal decline in the cultural affinity between Italians and EU2004/07 citizens after the enlargements took place of about one percent.

[36] Specifically, $R_{xx'}$ denotes the percentage marginal effect of the number of males of type x', $I_{x'}$, in the market on the number of single males of type x, $\mu(x,0)$, $R_{xx'} := \frac{1}{\mu(x,0)} \frac{\partial \mu(x,0)}{\partial I_{x'}}$.

cultural affinity and legal access, and we abstract from the other traits affecting the surplus, such as age or education. After the accession, the legal access coefficient β_{tp} is set to one as she becomes a European citizen, and only cultural differences matter. The highest surplus is then reached for a homogamous couple, followed by a couple with an Italian husband. The lowest surplus is achieved with a husband from an African country. Before the enlargements, the ordering of the surplus was different as it is higher when the husband is Italian. The surplus of a homogamous couple is only ranked second, and almost on par with a marriage including an EU citizen. This reversal of the ranking of the surplus, due to changes in the legal status, explains the large change in marriage markets during that period. Legal status incentives are sizeable enough to overturn homophily regarding nationality, but only within partners from European countries.

Panel B shows the surplus of a 27 years old woman from EU2004-07 countries, marrying either native or EU2004-07 men of different ages. If opting for a husband from the same country, the highest surplus is achieved with a slightly older man at 32 years of age. However, that man does not grant legal status to the couple in the period before the accession. In contrast, Italian husbands do, and the surplus of marrying such a man is higher. In fact, the surplus when marrying a native is higher for an age that ranges between 25 to 39. This represents a deviation of 7 years when marrying a younger or older man, which provides another metric to evaluate the importance of legal status.

Finally, Panel C shows the surplus of a high educated woman from an EU2004-07 country marrying either a native or a EU2004-7 man and with either a low or a high education. When marrying someone from the same country, the highest surplus is achieved when the husband has the same level of education. However, when marrying a native, the surplus is higher, even when the man is of lower education. Given that individuals with a low (high) education level have, on average 8 (14) years of education, the legal status premium represents about a 6 year difference in education within the couple.

5.3 Enlargements of the EU and welfare effects

The differences in marital choices induced by the EU enlargements likely generated heterogeneous changes in marital surplus. In this section, we evaluate the overall welfare effect of granting legal status to citizens of "new EU" countries and we measure the change in surplus for natives and foreigners. We consider the situation before the enlargements as our baseline. We then compute a counterfactual with a distribution of men and women identical to the baseline, but where citizens of "new EU" countries have a legal right to reside and work. In such a way, we isolate the legalization effect from a change in the composition of

men and women, potentially due to the policy itself and other factors affecting the Italian economy in the 2000s.[37]

Our model specifies a marital surplus, but not individual welfare, as, so far, we did not have to outline how the surplus is shared within the household.[38] As we want to compute the welfare separately for different types of individuals, we now take a stand on how this surplus is shared. Denote $\tilde{\Phi}_{Sp}(i,j)$ and $\omega_{Sp}(x_i)$ the marital surplus if man i is married to woman j and the value of being single in scenario $S = \{B, C\}$ - indicating the baseline or the counterfactual. We assume that the share of the surplus going to man i is pro-rata the value of staying single:

$$\tilde{\Phi}_{Sp}(i) = \tilde{\Phi}_{Sp}(i,j)\frac{\omega_{Sp}(x_i)}{\omega_{Sp}(x_i) + \omega_{Sp}(y_j)}, \quad S = B, C. \tag{11}$$

For a given individual, welfare may differ between the baseline and counterfactual for two reasons. First, there is a direct effect through a change in $\beta_{Sp}(x_i, y_j)$ affecting marital surplus. Second, there are general equilibrium effects as individuals married under the baseline may end up single in the counterfactual and vice-versa, and married individuals in the baseline may be married to someone different in the counterfactual with a different marital surplus and also different allocation weights. The evolution of individual welfare is therefore complex.

Figure 14 shows the results, for native or foreign men and women. The figure displays on the left the total change in welfare for each group of individuals and shows considerable heterogeneity. Native men, married in the baseline, experience a decline in welfare of about 0.6 percent. Married native women have a smaller decline of about 0.2 percent. Male citizens from the "new EU" countries experience a small decrease in their welfare, while foreign women experience a decrease of about 1.2 percent. We next display an Oaxaca decomposition of those welfare changes in four distinct categories. The first two are for married individuals, both in the baseline and counterfactual, and consist of a change in the marital surplus, holding allocation weights fixed, and a change in allocation weights, keeping the marital surplus fixed. The following two sets of results report the contribution to the welfare change of individuals moving from married to single and vice-versa. The figure reveals that the negative effect for native males mostly comes from a change in allocation weights and being single in the counterfactual. Indeed, native men are more likely to marry foreigners in the baseline and get a higher share of the surplus. In the counterfactual, they are more likely to remain single. This argument applies to some extent to native women as

[37] When we run counterfactuals, we fix $\varepsilon_t(i,j)$ as at baseline to isolate the effect of the policy.

[38] Identifying the sharing rule would require additional intra-household information.

well, although an increase in the marital surplus partly offsets the negative change in the allocation weights. For men from the "new EU" countries, the biggest welfare effect is from a higher marital surplus, as they are more likely to enter into homogamous marriages once they have legal status, which provides a higher marital surplus. They are also more likely to get married. Foreign women of those countries also experience an increase in welfare due to a higher likelihood of marriage.

Hence, general equilibrium effects induced by the enlargements had unintended redistributional consequences both within the market across different matches but also within a family across spouses. Overall, the policy redistributed marital surplus, with a loss for natives and a gain for foreign citizens, but also with a marked gender imbalance as foreign men benefit more from the policy change.

5.4 How does policy affect marriage markets?

We finally investigate two counterfactual policies. The first policy grants legal access to a broader set of individuals, while the second allows the inflow of a particular group of immigrants into the country.

5.4.1 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 immigrants. Similar policies have been enacted in France, Italy, and Spain. Their goal is to integrate irregular immigrants by granting legal residence and access to 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 assess the effect of generalized legalization on marriage markets using our estimated model and simulating a counterfactual where all immigrants in the economy are granted legal status, by assuming $\underline{\beta}_p = 1$ for everyone in the economy. Note that the variation we use to identify the model is closely related to such a policy. We solve the model with this new feature and compare it to the baseline. Figure 15 reports the change in marriage and fertility rates following the generalized access. The policy leads to a small increase in marriage rates in the whole economy (by about one percent) as well as a small increase in homogamous marriages of natives (of about 0.1 percent). These minor effects mask crucial changes in other markets as there is a sharp decrease in mixed marriages by about 7 percent,

mirrored by an even sharper increase in homogamous marriages among foreigners by about 22 percent. Moreover, while the policy has an overall small negative effect on fertility, the fertility rate increases in mixed marriages and in homogamous marriages among foreigners, in response to a change in the marital surplus of married couples.

While a legalization policy fosters the inclusion of foreigners in the labor market, it also has the unintended effect of reducing mixing in marriages while enhancing their selection with higher fertility. Given the strong intergenerational links in the transmission of cultural values within homogamous households, such a policy thus hamper the cultural integration of first generation immigrants with potentially severe and long-lasting consequences also on successive generations (Glazer et al., 1970; Suárez-Orozco and Suárez-Orozco, 2009; Schiller et al., 1995; Bisin and Tura, 2019).

5.4.2 Surge in African immigration

In recent years, there has been a marked increase in European countries of immigrants crossing through the Mediterranean. This surge has caused dissension among European countries and a debate on the long-term effect of such migration. While part of the argument revolves around the possible economic and cultural assimilation of those migrants, its impact on the marriage market is not prominent. However, there are reasons to believe marriage outcomes may lead to important issues. One such issue is the gender asymmetry in the marriage surplus and its consequences on marriage rates. Indeed, in Section 5.1, the marriage surplus in heterogamous marriages is larger when the native spouse is a man rather than a woman, possibly, leading to uneven marriage patterns for migrants by gender.

To test this hypothesis, we simulate the impact on the Italian marriage market of an increase in African immigrants. We assume that African immigrants are young (their age ranges between 22 and 32) and of low education, but balanced across sexes, which corresponds to recent migration experience. Figure 16 displays the results for new arriving migrants distinguishing between men and women, for the model baseline and with an increase of 7 and 15 percent in the size of that community. We report both marriage rates and the average age gap for Africans of age less than 32. We find that marriage rates are indeed different by gender: about 20 percent of foreign men get married as opposed to about 35 percent of women (top panels). Migrant women are more likely to marry natives than migrant men and are also more likely to marry within their own community. As the size of the African group increases, marriage rates decrease, mainly because of a decrease in marriages with natives. This is due to increasing competition for potential partners.

We also display the age gap for marriages formed by those migrants (bottom panels). As discussed above, there is a trade-off between the age and cultural gap between spouses.

African women marrying natives have a larger age gap, meaning their husband is six years older on average. When marrying within their own community, the age gap is only three years. For men, the age gap is much smaller in absolute value and negative, reflecting that those men marry women older than they are. This side of the market is heavily selected as fewer of those men marry and the women from their own community tend to get married to natives.

There is a large literature detailing the consequences of marriage on a range of outcomes such as earnings (marriage premium) but also crime and assimilation (Edlund et al., 2013). Hence, it is likely that the surge in immigration may result in a group of single foreign men facing difficulties in the labor market and, more generally, in integrating into the host country's society and economy.

6 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. Using both reduced-form techniques and a structural approach, we show that the accession of Eastern European countries profoundly changed the composition of mixed marriages. The migrants from the new European countries turned away from natives to marry more within their national communities. In turn, the natives changed their marriage pattern towards migrants who did not (yet) have legal access to the labor market. Therefore, our analysis stresses the importance of legal access to the labor market in marriage decisions and how it can overcome cultural differences. We estimate the value of legal rights to work and residency in terms of marital surplus to be as large as a 7 year gap between spouses or a 6 year gap in education. We also find that granting legal rights to migrants induces a transfer of marital welfare away from natives towards migrants with different redistributional effects along gender lines.

Hence, our analysis stresses the complex nature of legalization policies and evaluates their spillover effects on other important choices natives and migrants make. Those policies are meant to assimilate migrants through better inclusion in the labor market. We show that such policies decrease the incentives of targeted migrants to marry natives and therefore increase segregation in the marriage market while fostering selection inducing higher fertility. On the other side, though, a targeted policy such as the EU enlargements induces substitution patterns, allowing other immigrant groups further apart along cultural lines to have a higher possibility of intermarrying with natives.

References

Adda, J. and R. W. Cooper (2003). *Dynamic Economics: Quantitative Methods and Applications*. MIT press.

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

Alesina, A., Y. Algan, P. Cahuc, and P. Giuliano (2015). Family Values And The Regulation Of Labor. *Journal of the European Economic Association 13*(4), 599–630.

Alesina, A. and P. Giuliano (2010). The Power of the Family. *Journal of Economic Growth 15*, 93–125.

Alesina, A. and P. Giuliano (2011). Family Ties and Political Participation. *Journal of European Economic Association 9*(5), 817–839.

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 G. Tura (2019). Marriage, Fertility, and Cultural Integration in Italy. Working Paper 26303, National Bureau of Economic Research.

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.

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

Chiappori, P. and S. Oreffice (2008). Birth Control and Female Empowerment: An Equilibrium Analysis. *Journal of Political Economy 116*(1), 113–140.

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

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2015). The Becker–Coase Theorem Reconsidered. *Journal of Demographic Economics 81*(02), 157–177.

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

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

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

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

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

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

Ciscato, E., A. Galichon, and M. Gousse (2020). Like Attract Like? A Structural Comparison of Homogamy Across Same-Sex and Different-Sex Households. *Journal of Political Economy 128*(2), 740–781.

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.

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

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–96.

De Graaf, P. M. and M. Kalmijn (2001). Trends in The Intergenerational Transmission of Cultural and Economic Status. *Acta Sociologica 44*(1), 51–66.

Decker, C., E. H. Lieb, R. J. McCann, and B. K. Stephens (2013). Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market. *Journal of Economic Theory 148*, 778–792.

Devillanova, C., F. Fasani, and T. Frattini (2018). Employment of Undocumented Immigrants and the Prospect of Legal Status: Evidence from an Amnesty Program. *Industrial and Labor Relations Review 71*(4), 853–881.

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.

Edlund, L., H. Li, J. Yi, and J. Zhang (2013). Sex Ratios and Crime: Evidence from China. *The Review of Economics and Statistics 95*(5), 1520–1534.

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.

Fernandez, R. (2011). Does Culture Matter? In J. Benhabib, M. O. Jackson, and A. Bisin (Eds.), *Handbook of Social Economics*, Volume 1A. North-Holland.

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.

Figlio, D., P. Giuliano, U. Özek, and P. Sapienza (2019). Long-Term Orientation and Educational Performance. *American Economic Journal: Economic Policy 11*(4), 272–309.

Fisman, R., S. Iyengar, E. Kamenica, and I. Simonson (2006). Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment. *The Quarterly Journal of Economics 121*(2), 673–697.

Fisman, R., S. Iyengar, E. Kamenica, and I. Simonson (2008). Racial Preferences in Dating. *The Review of Economic Studies 1*(1), 117–132.

Fumagalli, S. and T. Boeri (2009). Inventory of migration policies (1990–2005): reforms and index on strictness of migration policy.

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

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

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

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

Glazer, N., D. P. Moynihan, et al. (1970). *Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians, and Irish of New York City*, Volume 63. MIT Press Cambridge.

Goodman-Bacon, A. (2021). Difference-in-differences with Variation in Treatment Timing. *Journal of Econometrics*.

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

Gualdani, C. and S. Sinha (2022). Partial identification in matching models for the marriage market. *Journal of Political Economy*.

Guiso, L., P. Sapienza, and L. Zingales (2003). People's Opium? Religion and Economic Attitudes. *Journal of Monetary Economics 50*(1), 225–282.

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.

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

Hofstede, G., G. J. Hofstede, and M. Minkov (2010). *Cultures and Organizations - Software of the Mind: Intercultural Cooperation and its Importance for Survival (3. ed.)*. McGraw-Hill.

Kalmijn, M. (1994). Assortative Mating by Cultural and Economic Occupational Status. *American Journal of Sociology 100*(2), 422–452.

Kalmijn, M. (1998). Intermarriage and Homogamy: Causes, Patterns, Trends. *Annual Review of Sociology 24*(1), 395–421.

Low, C. (2019). A "Reproductive Capital" Model of Marriage Market Matching. *Manuscript, Wharton School of Business*.

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.

Mourifié, I. (2019). A Marriage Matching Function with Flexible Spillover and Substitution Patterns. *Economic Theory 67*(2), 421–461.

Mourifié, I. and A. Siow (2021). The Cobb-Douglas Marriage Matching Function: Marriage Matching With Peer and Scale Effects. *Journal of Labor Economics 39*(S1), S239–S274.

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.

Schiller, N. G., L. Basch, and C. S. Blanc (1995). From Immigrant to Transmigrant: Theorizing Transnational Migration. *Anthropological quarterly*, 48–63.

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

Shephard, A. (2019). Marriage Market Dynamics, Gender, and the Age Gap. mimeo University of Pennsylvania.

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

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

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

Suárez-Orozco, C. and M. M. Suárez-Orozco (2009). *Children of Immigration*. Harvard University Press.

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

Vickery, A. and D. Anderberg (2021). The Role of Own-Group Density and Local Social Norms for Ethnic Marital Sorting: Evidence from the UK. *European Economic Review 138*, 103774.

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

Figures and Tables

Figure 1: Intermarriage rate across European countries, 2015

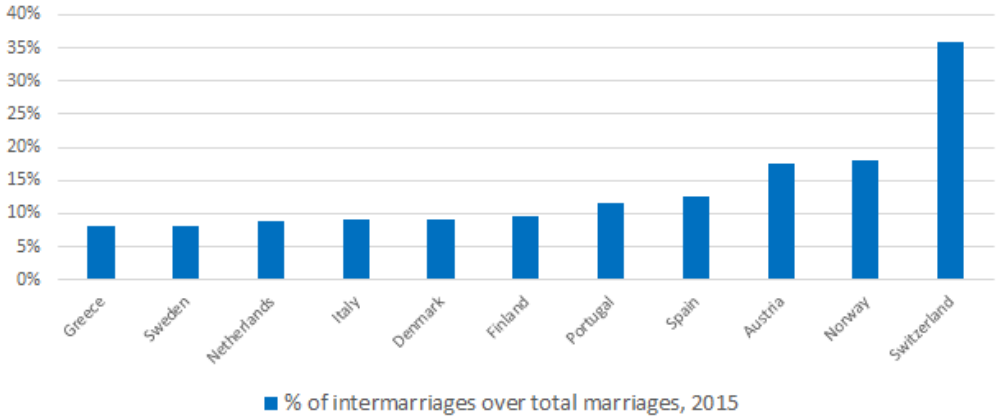


Figure 2: Number of foreign residents in Italy by area of origin (millions), 1994-2017

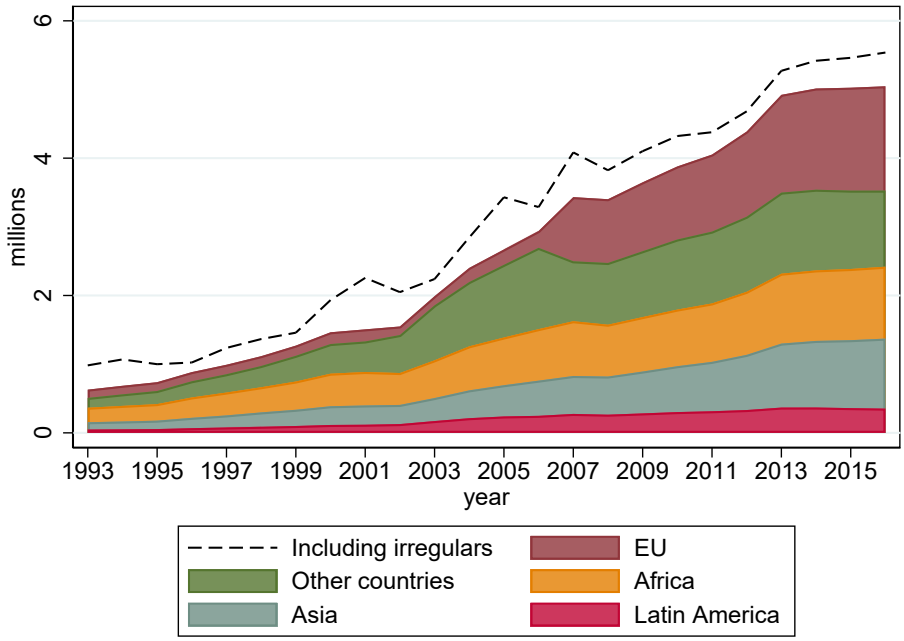
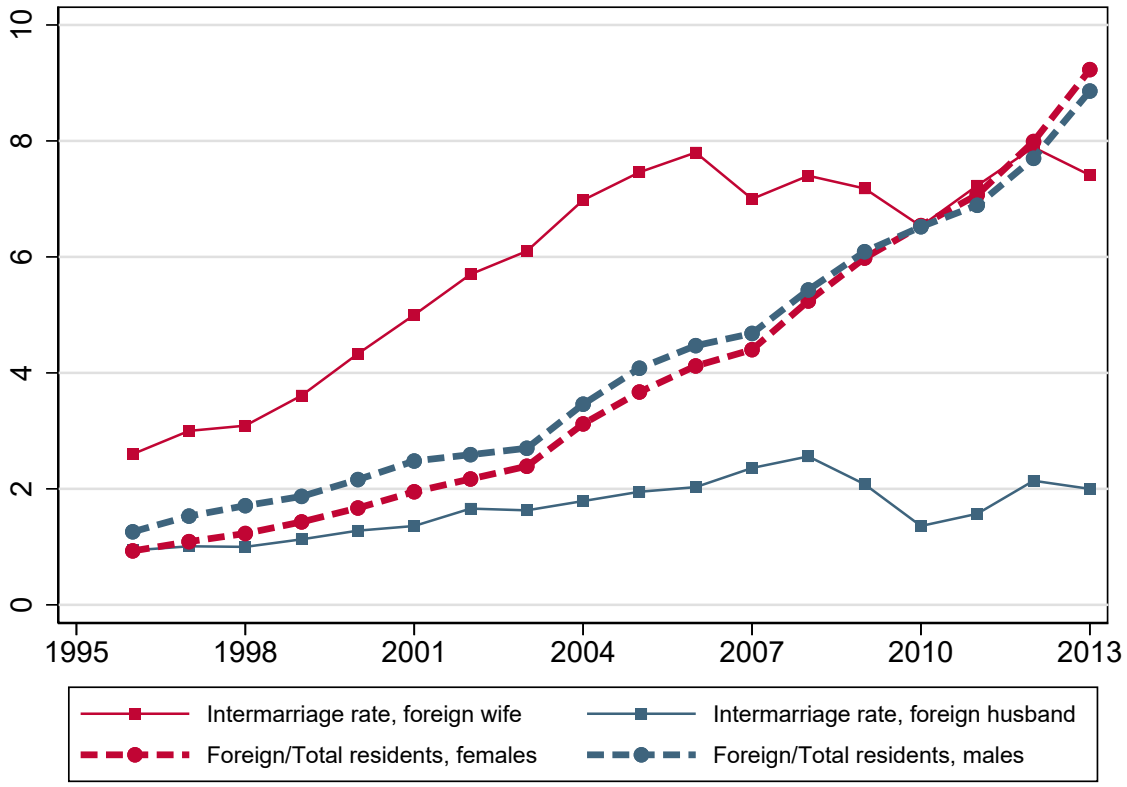


Figure 3: Intermarriage rate and ratio of foreign residents over total population, 1996-2013



Notes: 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 4: Timeline of the 2004 and 2007 EU enlargements

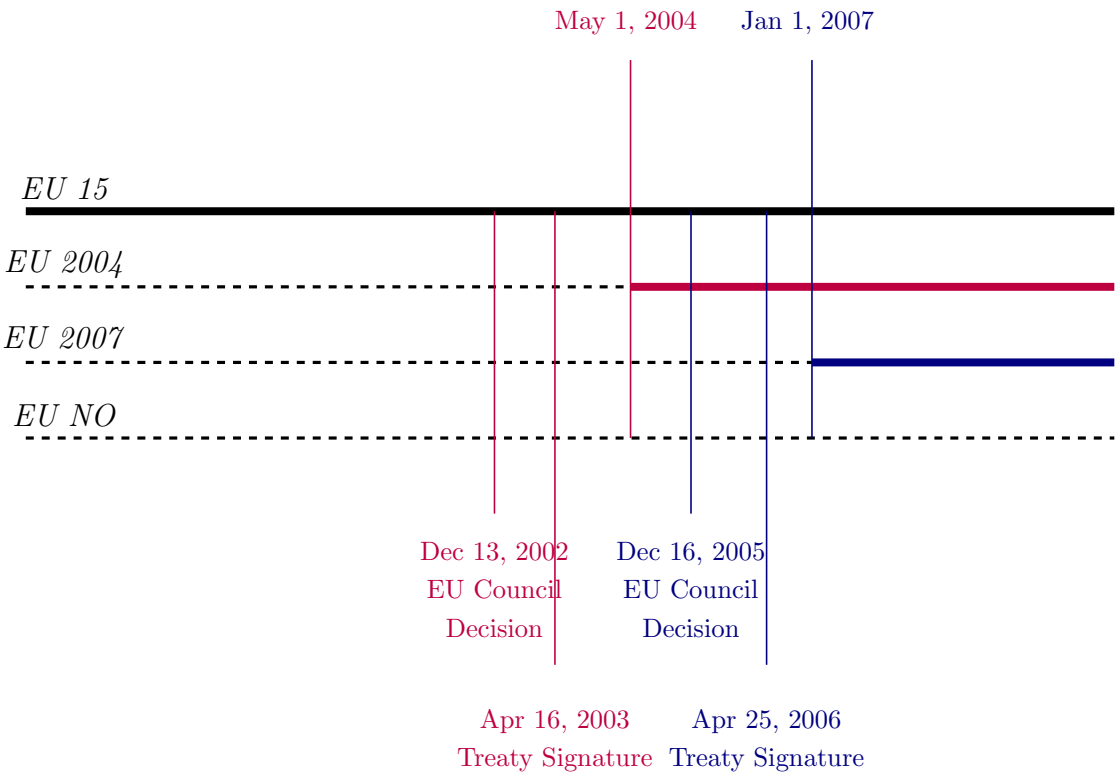
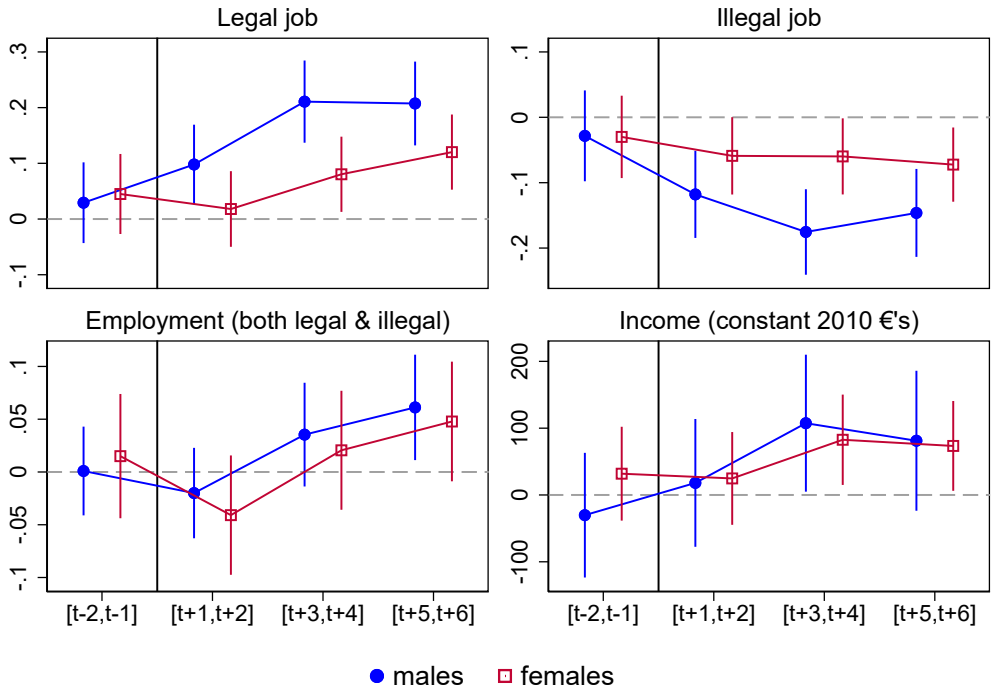


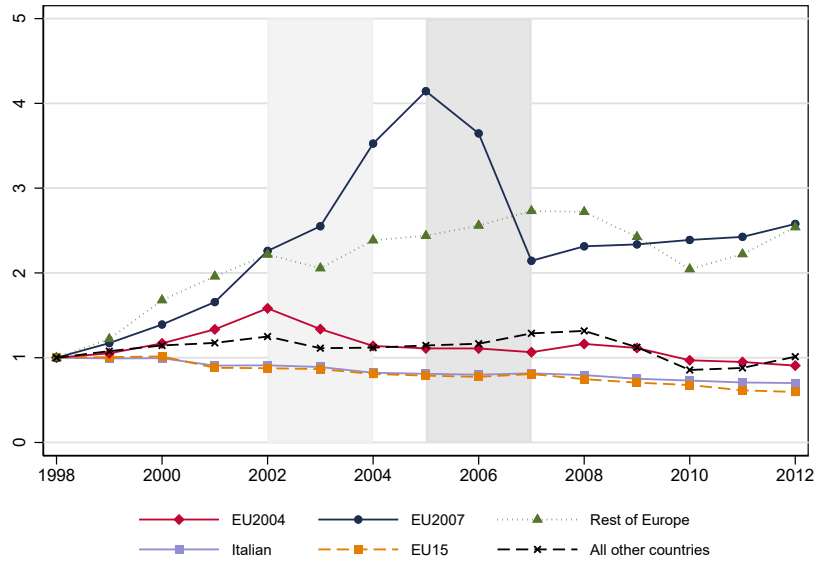
Figure 5: The effect of the EU enlargements on the labor market outcomes of new EU citizens



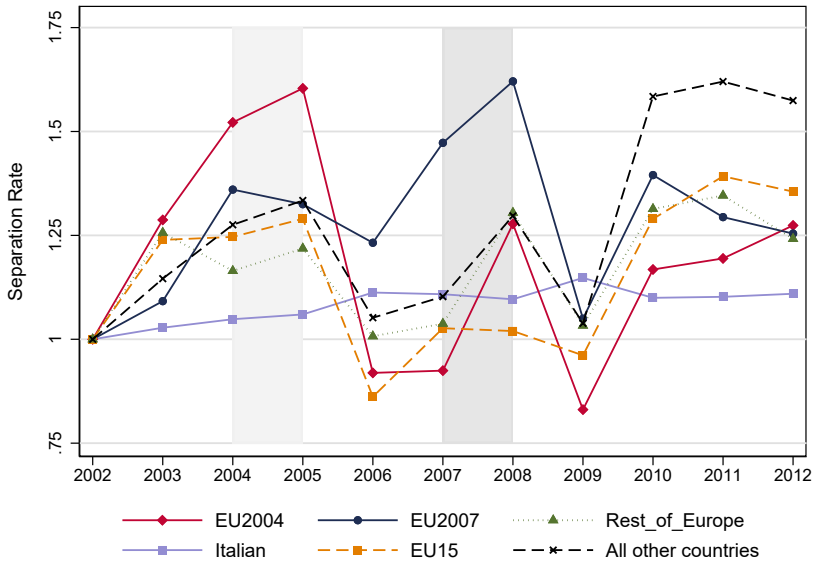
Notes: These graphs show the estimated dynamic treatment effects of admission to the EU on labor market outcomes, estimated using an event-study approach on individual-level data from the ISMU survey (described in Section 2.3 and Appendix D). The sample stacks two sub-samples including all new EU citizens interviewed between two years before and six years after they entered the EU, and all immigrants from other countries surveyed during the same period. The dependent variables are indicated on top of each graph and specifications include country of birth fixed effects, year fixed effects, and a vector of individual characteristics (a quadratic polynomial in age, education, and marital status). Year fixed effects and individual characteristics are always interacted with dummy indicators for each stacked sub-sample. 95% confidence intervals are also shown in the graphs. The vertical line in each graph denotes the moment of admission to the EU.

Figure 6: 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)



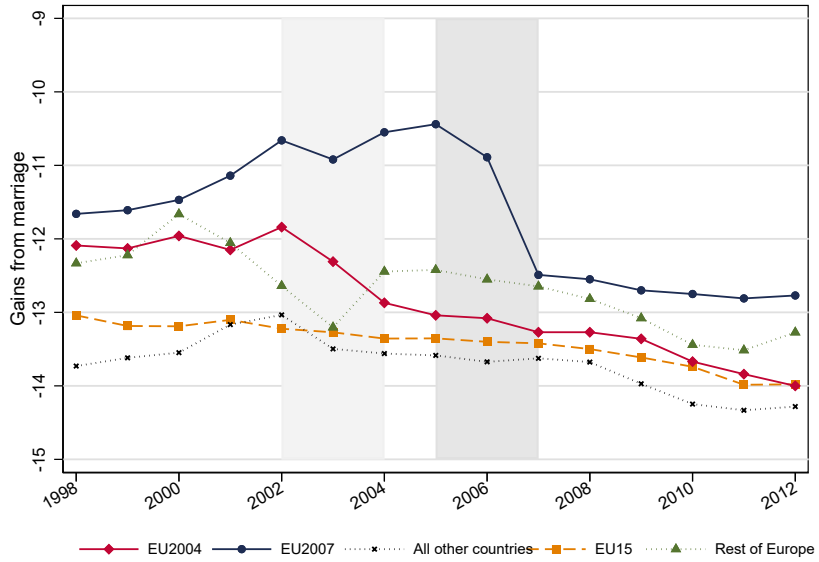
(b) Separation rates, marriages formed since 1998 (index=1 in 2002)



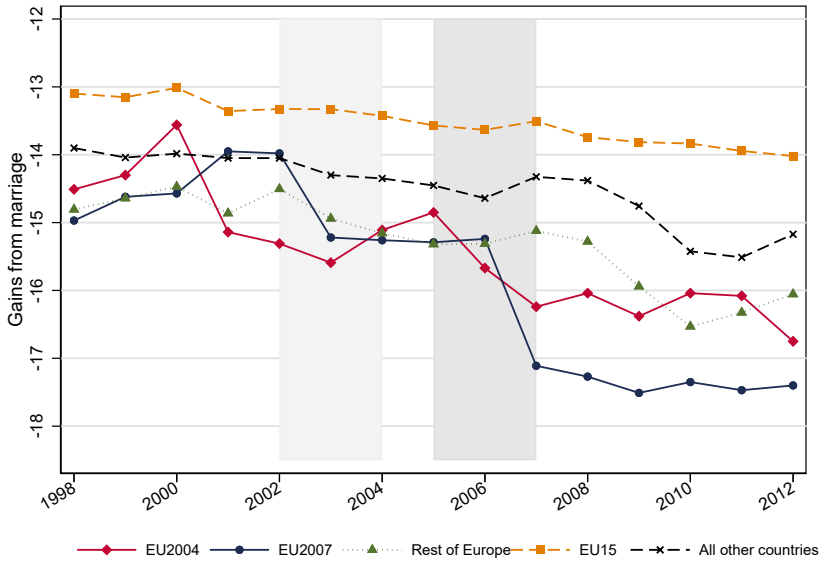
Notes: Panel (a) shows the share of marriages in which (at least) one spouse was Italian, by area of origin of the other spouse. The series is standardized to 1 in 1998. For these same marriages, Panel (b) shows the share of separations by area of origin. The series of separations is standardized to 1 in 2002. The classification of countries is reported in Table A1.

Figure 7: Gains from marriage, heterogamous couples formed by natives and immigrants, by area of origin of the foreign spouse

(a) Native husband – foreign wife



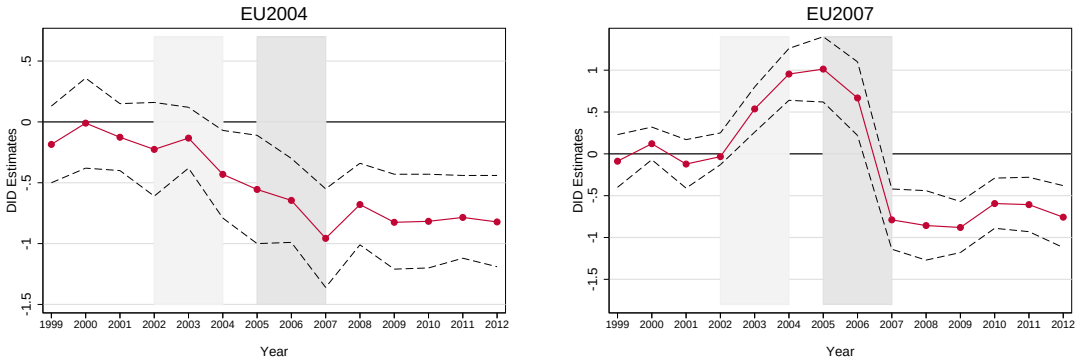
(b) Native wife – foreign husband



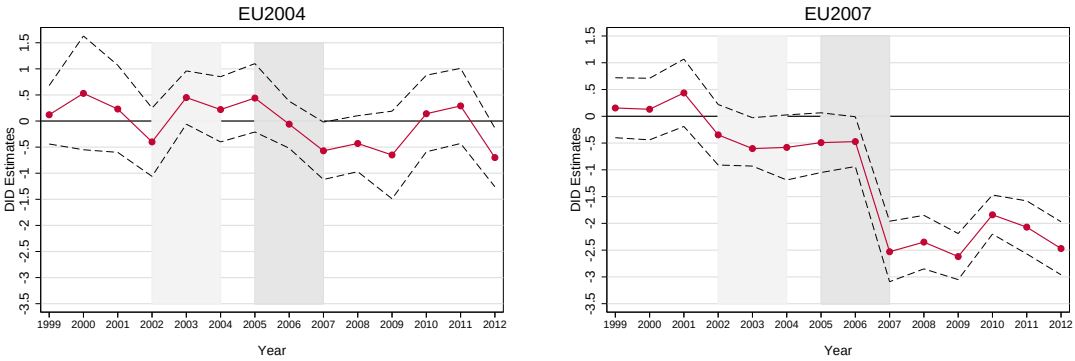
Notes: The graphs plot the gains from marriage over the period 1998-2012 by immigrants' area of origin and gender of the foreign spouse. The shaded areas denote the periods between the announcement and implementation of the EU enlargements. Gains from marriage are measured as in equation (1). The classification of countries is reported in Table A1. *Source*: ISTAT, marriage records from vital statistics registries (1998-2012) and individual Census data.

Figure 8: Changes in gains from marriage between natives and new EU citizens, year-specific estimates

(a) Native husband – foreign wife

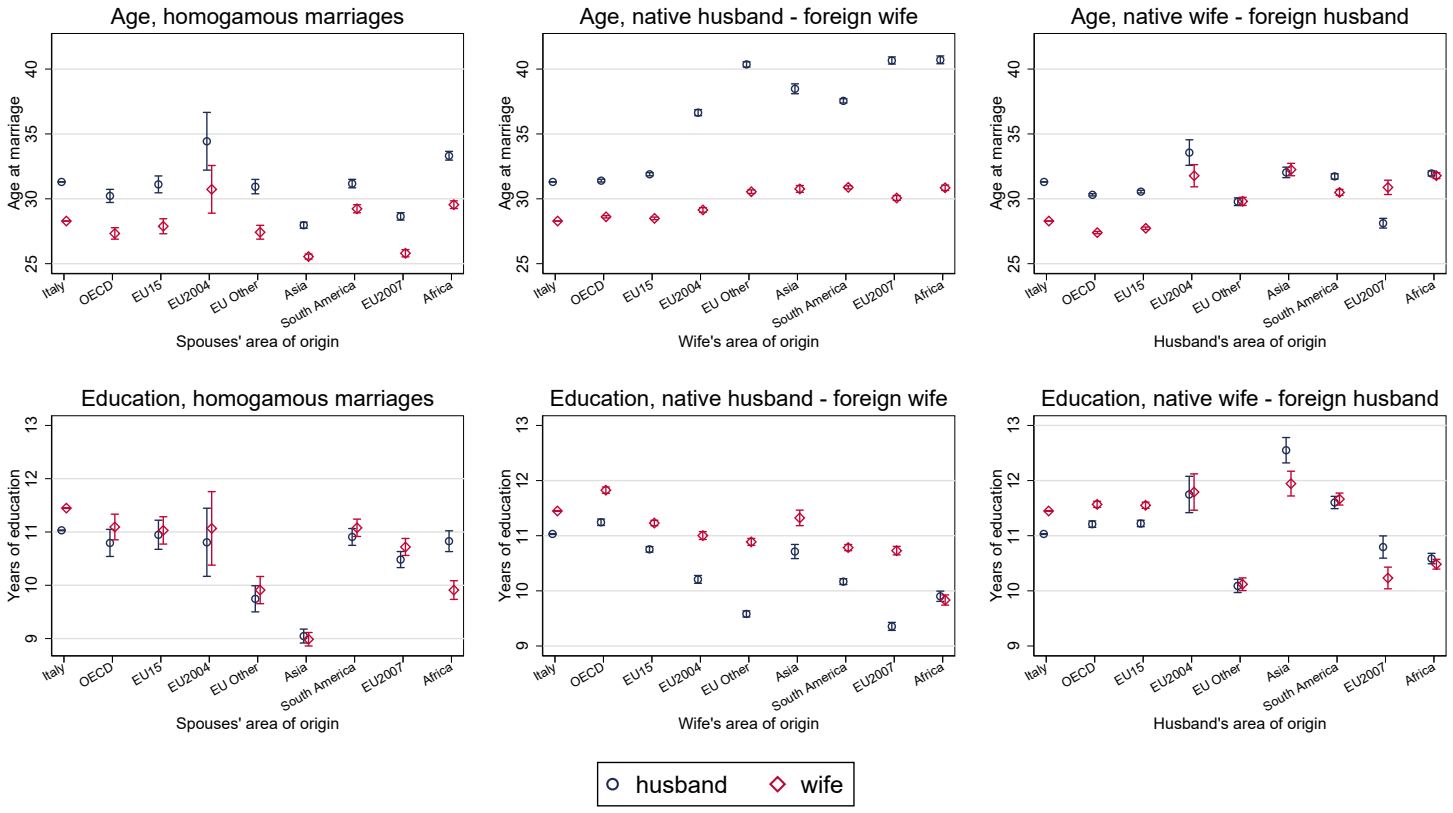


(b) Native wife – foreign husband



Notes: 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 ($EU2004$ and $EU2007$) 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 9: Spouses' characteristics in homogamous marriages and intermarriages, before enlargement periods (1998-2002)



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

Figure 10: Marital surplus related to legal access and cultural affinity and its policy variation

Before enlargements:

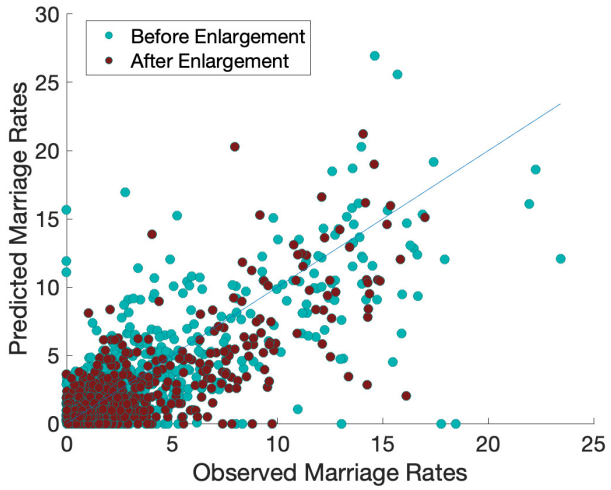
		Women: Italian	Women: New EU	Women: Other
Men	Italian	1	$\bar{\beta}^{\male}\alpha_{I/N}$	$\bar{\beta}^{\male}\alpha_{I/O}$
Men	New EU	$\bar{\beta}^{\female}\alpha_{N/I}$	$\underline{\beta}$	$\underline{\beta}\alpha_{N/O}$
Men	Other	$\bar{\beta}^{\female}\alpha_{O/I}$	$\underline{\beta}\alpha_{O/N}$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

After enlargements:

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

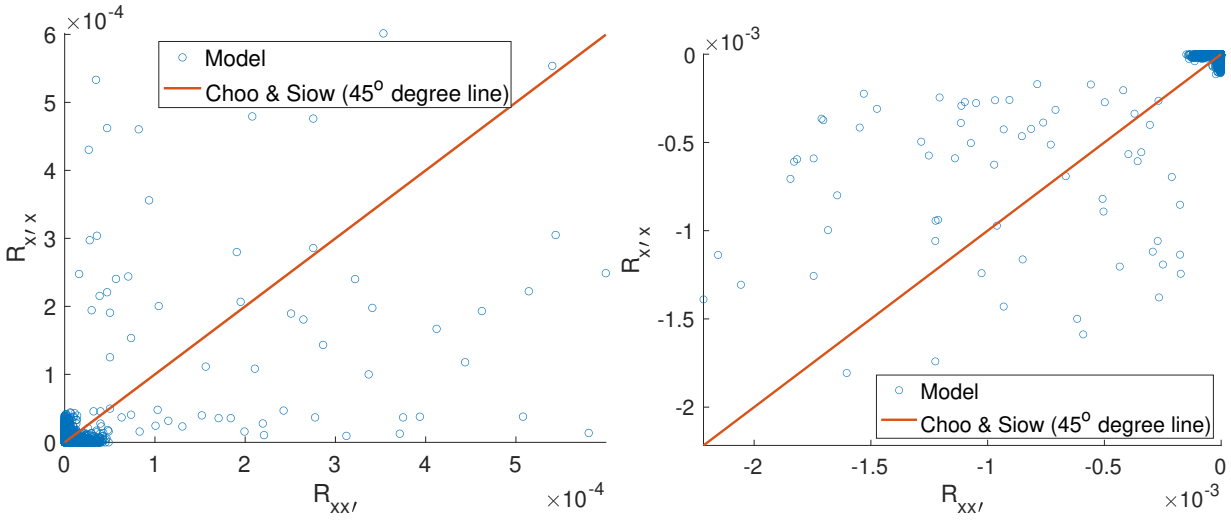
Notes: Each cell shows the marital surplus associated with legal access to work and with culture for spouses of different origins. $\alpha_{x/y}$ are cultural affinity parameters affecting marriage surplus, $\underline{\beta}$ is the penalty for no legal access, $\bar{\beta}^{\female}$ and $\bar{\beta}^{\male}$ 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.
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 11: Fit of model across marriage markets



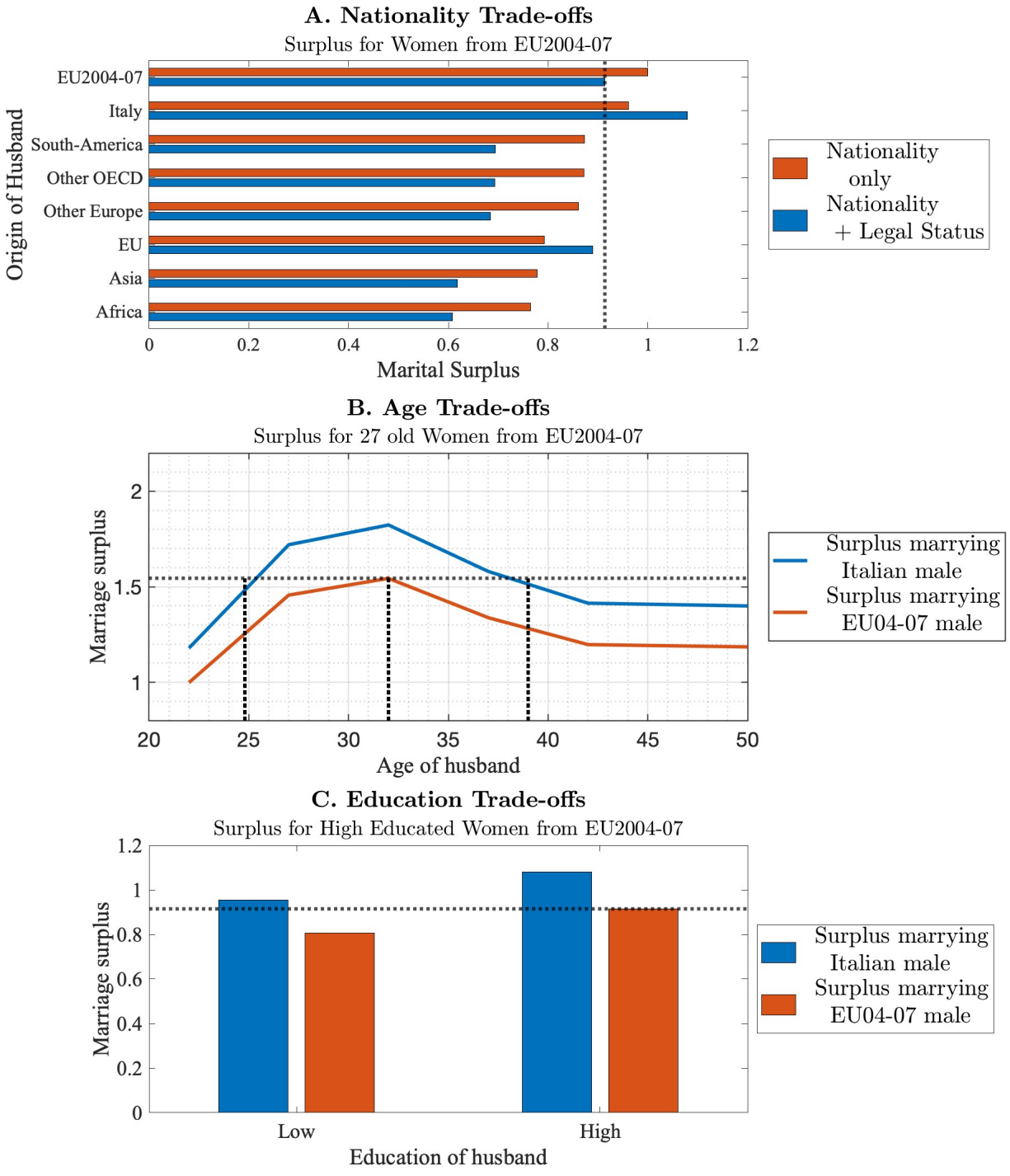
Notes: Each dot represents a particular marriage market, defined by nationality, age, education and province of residence of the spouses, before (blue) and after (black) the EU enlargements. Marriage rates have been multiplied by 1000 for ease of reading.

Figure 12: Substitution elasticities



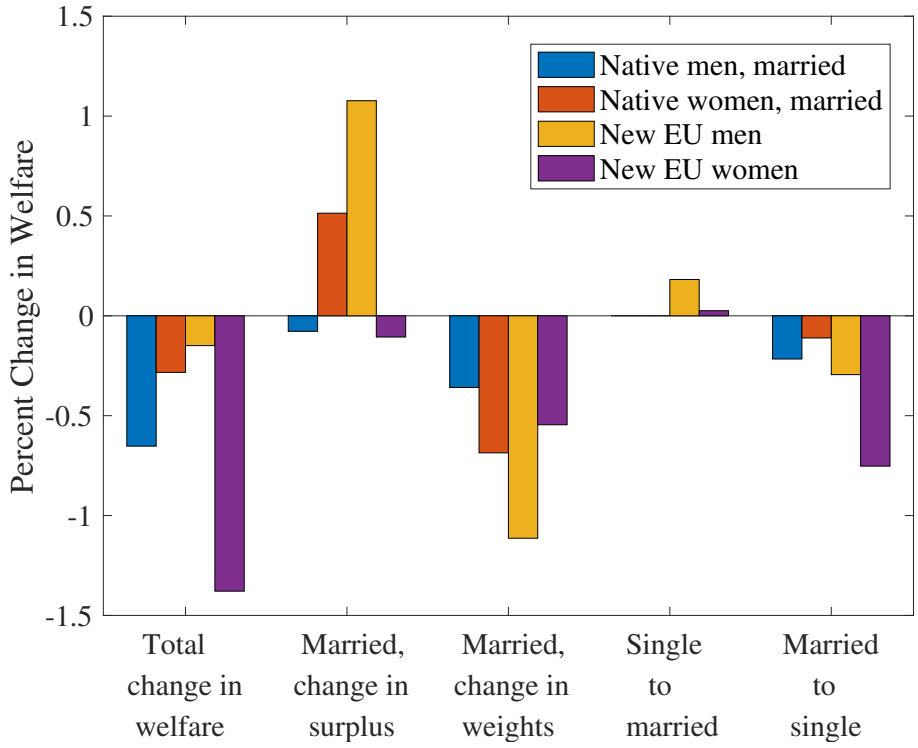
Notes: The Figures plot the marginal effects $R_{xx'}$ as a function of the marginal effects $R_{x'x}$, where $R_{xx'}$ is the marginal effect of the number of males of characteristics x on the log number of singles (male or female) of characteristics x'. The Choo and Siow model imposes i) that $R_{xx'} = R_{x'x}$ which is depicted by the 45 degree line (red line), ii) that $R_{xx'} > 0$ for the effects on single men and $R_{xx'} < 0$ for single women.

Figure 13: Trade-offs: Legal status and other characteristics



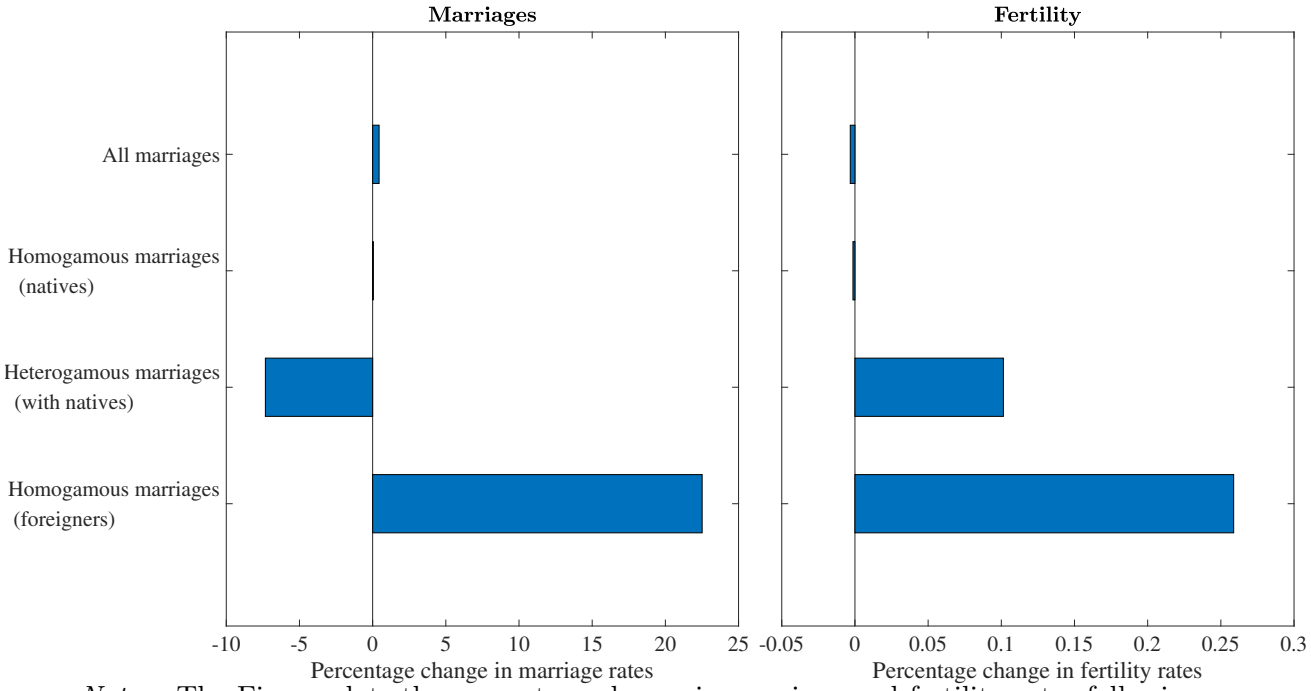
Notes: The Figure plots the surplus from the perspective of a woman from an EU2004-07 country marrying different types of men. It depicts the surplus related to legal status and either nationality (panel A), age (panel B) and education (panel C).

Figure 14: Change in surplus following the EU enlargements



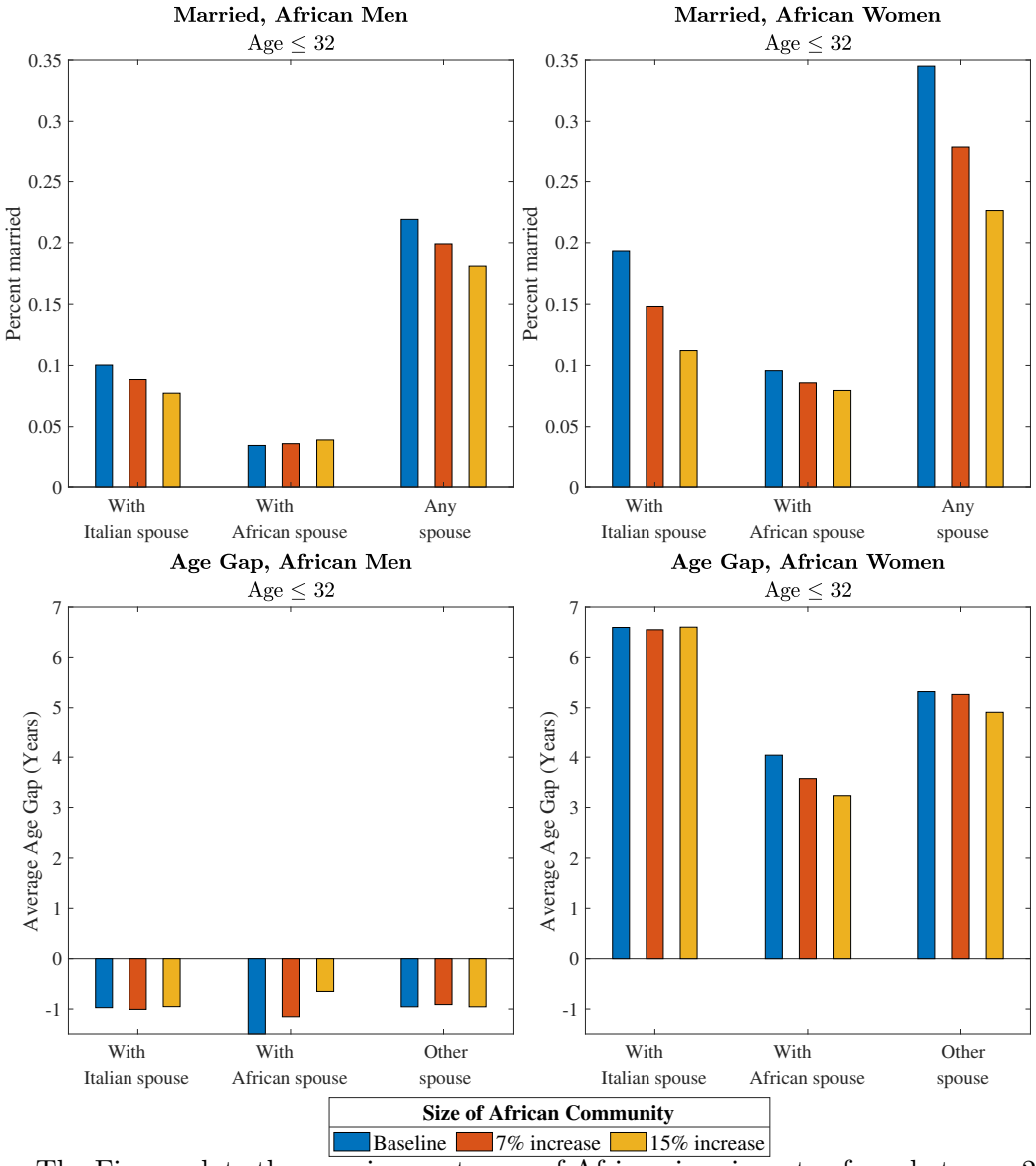
Notes: The Figure plots the percentage change in the surplus between a baseline scenario where citizens of new EU countries are not granted legal rights to work and a counterfactual scenario where they have legal rights. The figure displays the total change in surplus and a Oaxaca decomposition which is composed of four distinct components, summing up to the total change. The first two components are for individuals married in both baseline and counterfactual scenarios with, respectively, the change in allocation weights keeping surplus fixed and the change in the surplus keeping allocation weights constant; see equation (11). The other two components are the contributions to the welfare of individuals going from married to single and vice-versa.

Figure 15: Change in marriage and fertility rates following a generalized legal access



Notes: The Figure plots the percentage change in marriage and fertility rates following a generalized policy granting legal status to all foreigners.

Figure 16: Marriage patterns of additional African immigrants



Notes: The Figure plots the marriage outcome of African immigrants of age between 22-32 and of low education for various sizes of the African community. The age gap is measured as the average of the age of the husband minus the one of the wife.

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

	Males and females			**Males**			**Females**		
	regular	irregular	diff.	regular	irregular	diff.	regular	irregular	diff.
				Panel A: Labor market outcomes					
Employed	0.782	0.746	0.036	0.886	0.731	0.156	0.644	0.777	-0.133
	(0.001)	(0.004)	(0.004)	(0.001)	(0.005)	(0.004)	(0.002)	(0.007)	(0.008)
Legal job	0.699	0.062	0.638	0.813	0.059	0.754	0.550	0.067	0.483
	(0.002)	(0.002)	(0.004)	(0.002)	(0.003)	(0.005)	(0.003)	(0.004)	(0.008)
Illegal job	0.082	0.684	-0.602	0.073	0.672	-0.598	0.094	0.710	-0.616
	(0.001)	(0.004)	(0.003)	(0.001)	(0.005)	(0.004)	(0.001)	(0.007)	(0.005)
Income	798.0	562.3	235.7	998.9	553.3	445.6	533.7	580.9	-47.2
	(2.3)	(5.0)	(6.5)	(3.1)	(6.2)	(8.2)	(2.8)	(8.4)	(9.3)
				Panel B: Other individual characteristics					
Male	0.568	0.672	-0.104						
	(0.002)	(0.004)	(0.005)						
Age	35.4	31.4	4.0	35.6	30.2	5.4	35.3	33.8	1.4
	0.03	0.07	0.08	0.04	0.08	0.10	0.04	0.15	0.15
Married	0.628	0.320	0.308	0.625	0.286	0.338	0.633	0.389	0.244
	(0.002)	(0.004)	(0.005)	(0.002)	(0.005)	(0.006)	(0.002)	(0.008)	(0.008)
N. of children	1.310	0.761	0.549	1.271	0.596	0.675	1.362	1.099	0.263
	(0.004)	(0.011)	(0.013)	(0.006)	(0.013)	(0.017)	(0.006)	(0.020)	(0.021)
High school	0.410	0.395	0.015	0.395	0.367	0.028	0.429	0.452	-0.023
	(0.002)	(0.005)	(0.005)	(0.002)	(0.005)	(0.006)	(0.003)	(0.008)	(0.008)
College	0.148	0.104	0.043	0.124	0.085	0.039	0.179	0.144	0.035
	(0.001)	(0.003)	(0.003)	(0.001)	(0.003)	(0.004)	(0.002)	(0.006)	(0.006)

Note: This table compares average labor market outcomes (Panel A) and other individual characteristics (Panel B) between regular and irregular immigrants responding to the ISMU survey. Standard errors are reported in parentheses.

Table 2: 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	482,601	188,169	294,432	28,876	11,263	17,613
NewEU	100,340	17,702	82,638	5,886	865	5,021
All	7,317,582	3,658,791	3,658,791	416,124	208,062	208,062
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	31.3	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 3: Gains from marriage before and after the EU enlargements, difference-in-differences estimates

	native male - foreign female			*native female - foreign male*		
	(1)	(2)	(3)	(4)	(5)	(6)
$newEU \times PostEU$	-0.767***	-0.862***	-1.058***	-1.249***	-1.027***	-1.182***
	(0.053)	(0.059)	(0.065)	(0.123)	(0.094)	(0.098)
Observations	50,349	50,349	50,341	31,125	31,125	31,119
N. Marriages	3,402,324	3,402,324	3,402,312	3,239,135	3,239,135	3,239,127
R-squared	0.645	0.809	0.811	0.678	0.845	0.840
Year, prov, country FE	Yes	Yes	Yes	Yes	Yes	Yes
Country x trend	No	Yes	Yes	No	Yes	Yes
Province x trend	No	Yes	Yes	No	Yes	Yes
Province x country FE	No	Yes	Yes	No	Yes	Yes
Include irregular singles	Yes	Yes	No	Yes	Yes	No

Note: 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)-(3) and (4)-(6) 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 foreign spouse's country of origin. 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 specifications include country of origin, province, and year fixed effects; specifications in columns (2-3) and (5-6) include, in addition, linear trends interacted with province and country of origin fixed effects as well as province × country of origin fixed effects. In columns (3) and (6), the dependent variable is computed excluding (imputed) irregular immigrants at the denominator of the gains from marriage in equation (1). Regressions 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 enlargements, difference-in-differences estimates (robustness to controlling for the number of singles)

	native male (IT) - foreign female (y)				native female (IT) - foreign male (x)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$newEU \times PostEU$	-0.857***	-0.955***	-0.936***	-0.983***	-0.878***	-1.038***	-0.916***	-1.017***
	(0.066)	(0.084)	(0.086)	(0.092)	(0.105)	(0.090)	(0.108)	(0.141)
$\ln(SingleMales)$	-0.991***		-1.000***		-0.808***		-0.842***	
	(0.012)		(0.026)		(0.040)		(0.041)	
$\ln(SingleFemales)$	-0.749***		-0.750***		-1.151***		-1.189***	
	(0.032)		(0.029)		(0.132)		(0.142)	
$\ln(SingleMales_y)$		-0.121***	0.089***					
		(0.037)	(0.027)					
$\ln(SingleFemales_x)$						-0.158***	0.180***	
						(0.056)	(0.022)	
Observations	50,439	44,721	44,721	44,721	31,125	29,539	29,539	29,539
R-Squared	0.826	0.817	0.831	0.837	0.860	0.848	0.861	0.866
Interacted FEs & trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nationality x Singles	No	No	No	Yes	No	No	No	Yes

Note: 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 foreign spouse's country of origin. 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 year fixed effects, linear trends interacted with province and country of origin fixed effects, and province × country of origin fixed effects. In addition, columns (1) and (3) includes the logarithm of the number of single females in each cell; columns (2) and (3) includes the logarithm of the number of single males from country y; and column (4) includes all these variables interacted with country of origin fixed effects. Analogously, columns (5) and (7) includes the logarithm of the number of single males in each cell; columns (6) and (7) includes the logarithm of the number of single females from country x; and column (8) includes all these variables interacted with country of origin fixed effects. All regressions 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: Gains from marriage and cohabitation before and after the EU enlargements, difference-in-differences estimates

	native male - foreign female			*native female - foreign male*		
	marriages	cohabitations	both	marriages	cohabitations	both
	(1)	(2)	(3)	(4)	(5)	(6)
$newEU \times PostEU$	-1.290***	0.575***	-0.872***	-1.554***	-0.543	-1.316***
	(0.118)	(0.191)	(0.086)	(0.362)	(0.383)	(0.291)
Observations	11,758	8,740	12,360	8,499	5,421	9,118
R-squared	0.909	0.905	0.910	0.900	0.928	0.899
Interacted FEs & trends	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table shows the effect of admission to the EU on gains from marriage and cohabitation, estimated using the difference-in-differences specification (2). The sample includes all heterosexual marriages (columns 1 and 4), cohabitations outside a legal relationship (columns 2 and 5) and the sum of the marriages and cohabitations (columns 3 and 6) formed between 1998 and 2012 in which at least one of the spouses is Italian. Columns (1)-(3) and (4)-(6) 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 foreign spouse's country of origin. 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 specifications include year fixed effects, linear trends interacted with province and country of origin fixed effects, and province $\times$ country of origin fixed effects. Regressions 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 6: Hazard rate of separations before and after the EU enlargements, 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 post enlarg.	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 post enlarg.	-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 pre enlarg.		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 post enlarg.		-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

Note: 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 enlargements; the subsequent years; couples formed before the enlargements; and couples formed after the enlargements 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 7: Probability of having a child before and after the EU enlargements, difference-in-differences estimates

	native male - foreign female				*native female - foreign male*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$newEU \times PostEU$	0.087***	0.083***	0.087***	0.105***	0.055***	0.095***	0.075***	0.047***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.018)	(0.022)	(0.021)	(0.023)
Observations	3169914	3169914	3150633	3150633	3011556	3011556	3010393	3010393
R-squared	0.052	0.061	0.050	0.108	0.036	0.045	0.036	0.094
Prov, Year, CO FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prov Trends	No	Yes	No	Yes	No	Yes	No	Yes
CO Trends	No	Yes	No	Yes	No	Yes	No	Yes
Sample	all years	all years	restricted	restricted	all years	all years	restricted	restricted

Note: This table shows the effect of admission to the EU on fertility choices, estimated using a difference-in-differences specification. 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 the probability of having a child within three years of 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 specifications include country of origin, province, and year fixed effects. Standard errors clustered at the province level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Fertility and separation, observed and predicted moments

A. Fertility

	Observed	Predicted
Native male - foreign female, DiD	0.105(0.007)	0.10579
Native male - foreign female, level	0.527(0.026)	0.51934
Native female - foreign male, DiD	0.0471(0.023)	0.048173
Native female - foreign male, level	0.527(0.026)	0.51545

B. Separation

	Before enlargements		During enlargements		After enlargements	
	Observed	Predicted	Observed	Predicted	Observed	Predicted
Italian-New EU	0.0071 (0.0022)	0.0075	0.0098 (0.0025)	0.0231	-0.0051 (0.001)	-0.0011
Italian-Others	0.0018 (0.0007)	0.0119	0.0039 (0.0007)	0.0118	0.0044 (0.0008)	0.0079
Non Italian - Non Italian	-0.0047 (0.0017)	-0.0177	-0.0036 (0.001)	-0.0259	-0.0034 (0.001)	-0.0299
Woman >10 year younger	0.0011 (0.0005)	0.0179	0.0019 (0.0004)	0.0175	0.0022 (0.0004)	-0.0046
Woman >10 year older	0.0029 (0.002)	0.0069	0.0099 (0.0022)	0.008	0.0142 (0.0027)	0.0031
Same education	0.0001 (0.0004)	-0.002	0.0011 (0.0003)	-0.002	0.0008 (0.0002)	-0.0091
Constant	0.0122 (0.0003)	0.0197	0.0108 (0.0002)	0.0198	0.0107 (0.0002)	0.0299

Note: In this Table, panel A displays the observed and predicted moments related to the probability of having a child, as reported in Table 7. Panel B displays the auxiliary coefficients in a linear regression of an indicator variable equal to one if a couple separates within 3-4 years after marriages. Three separate regressions are run for couples married and followed before the enlargements, for couples married before and followed through the enlargements and couples married after the enlargements. The regression is defined in equation (10). Standard errors are reported in parentheses.

Table 9: Estimated parameters: surplus linked to legal access and cultural affinity

Parameters	Estimates	
Panel A: Legal access		
No legal access ($\underline{\beta}$), No shadow economy	0.914	(0.004)
No legal access ($\underline{\beta}$), shadow economy 100%	0.971	(0.004)
Legal access, native husband - foreign wife ($\bar{\beta}^{\male}$), No shadow economy	1.12	(0.004)
Legal access, native wife - foreign husband ($\bar{\beta}^{\female}$), No shadow economy	1.15	(0.003)
Legal access, native husband - foreign wife ($\bar{\beta}^{\male}$), shadow economy 100%	1.1	(0.011)
Legal access, native wife - foreign husband ($\bar{\beta}^{\female}$), shadow economy 100%	1.13	(0.003)
Panel B: Outside value		
Average outside value, before enlargement, North	6.29	(0.002)
Outside value, before enlargement, South	6.3	(0.003)
Outside value, after enlargement, North	6.45	(0.002)
Outside value, after enlargement, South	6.46	(0.003)

Panel C: Cultural affinity

		Women							
		Italian	EU15	New EU	Other Europe	Africa	Asia	South America	OECD
Men	Italian	1	0.99	0.96	0.91	0.9	0.86	0.9	0.96
		-	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)	(0.001)
	EU	0.99	1	0.79	0.75	0.8	0.66	0.73	0.81
		(0.001)	-	(0.005)	(0.005)	(0.001)	(0.003)	(0.001)	(0.001)
	New EU	0.92	0.79	1	0.75	0.67	0.68	0.76	0.76
		(0.001)	(0.005)	-	(0.005)	(0.004)	(0.001)	(0.007)	(0.001)
	Oth. Europe	0.81	0.75	0.75	1	0.8	0.85	0.74	0.73
		(0.001)	(0.005)	(0.005)	-	(0.002)	(0.001)	(0.001)	(0.001)
	Africa	0.8	0.8	0.67	0.8	1	0.56	0.82	0.88
		(0.005)	(0.001)	(0.004)	(0.002)	-	(0.001)	(0.001)	(0.001)
	Asia	0.73	0.66	0.68	0.85	0.56	1	0.84	0.92
		(0.004)	(0.003)	(0.001)	(0.001)	(0.001)	-	(0.001)	(0.001)
	S America	0.8	0.73	0.76	0.74	0.82	0.84	1	0.87
		(0.001)	(0.001)	(0.007)	(0.001)	(0.001)	(0.001)	-	(0.001)
	OECD	0.92	0.81	0.76	0.73	0.88	0.92	0.87	1
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	-

Note: The table displays in Panel A the part of the surplus associated with legal access obtained through marriage. Panel B displays the outside value of remaining single. Panel C. reports the part of the surplus associated with the country of origin; see equation (6) and Figure 10 for a description. The estimation relies on the observation of 1.4 million marriages.

ONLINE APPENDIX

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

A Additional figures and tables

Table A1: 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 A2: The effect of the EU enlargements on labor market outcomes of new EU citizens.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Employed		Legal job		Illegal job		Income	
Panel A: Males								
newEU x PostEU	0.018	0.016	0.108***	0.105***	-0.090***	-0.089***	85.264**	82.588**
	(0.016)	(0.016)	(0.026)	(0.025)	(0.023)	(0.023)	(33.347)	(32.525)
Constant	0.882***		0.685***		0.198***		812.380***	
	(0.005)		(0.007)		(0.006)		(8.389)	
Observations	76,865	76,865	76,865	76,865	76,865	76,865	76,865	76,865
Controls	NO	YES	NO	YES	NO	YES	NO	YES
R-squared	0.038	0.071	0.036	0.111	0.029	0.066	0.048	0.104
Panel B: Females								
newEU x PostEU	-0.013	-0.007	0.013	0.018	-0.026	-0.025	13.854	20.291
	(0.021)	(0.021)	(0.025)	(0.025)	(0.021)	(0.021)	(24.696)	(24.265)
Constant	0.637***		0.427***		0.211***		452.250***	
	(0.010)		(0.011)		(0.008)		(10.494)	
Observations	53,621	53,621	53,621	53,621	53,621	53,621	53,621	53,621
Controls	NO	YES	NO	YES	NO	YES	NO	YES
R-squared	0.159	0.216	0.082	0.129	0.049	0.057	0.109	0.161

Note: This table shows the effect of admission to the EU on labor market outcomes, estimated using an event-study approach on individual-level data from the ISMU survey (described in Section 2.3 and Appendix D). The sample stacks two sub-samples including, respectively, all $EU2004$ and $EU2007$ immigrants interviewed between two years before and six years after they entered the EU together with all immigrants from other countries surveyed during the same period. The dependent variables are indicated on top of each column and the main explanatory variable is the interaction between being a citizen $EU2004$ or $EU2007$ countries and a dummy for the period after each enlargement. All specifications include, in addition, country of birth and year fixed effects, and specifications in even columns also include a vector of individual characteristics (a quadratic polynomial in age, education, and marital status). Year fixed effects and individual characteristics are always interacted with dummy indicators for each stacked sub-sample. Standard errors are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Gains from marriage before and after the EU enlargements, difference-in-differences estimates (non-weighted)

	native male - foreign female			*native female - foreign male*		
	(1)	(2)	(3)	(4)	(5)	(6)
$newEU \times PostEU$	-0.676***	-0.862***	-1.058***	-1.249***	-1.027***	-1.182***
	(0.053)	(0.059)	(0.065)	(0.123)	(0.094)	(0.098)
Observations	50,349	50,349	50,341	31,125	31,125	31,119
R-squared	0.645	0.809	0.811	0.678	0.845	0.840
Year, prov, country FE	Yes	Yes	Yes	Yes	Yes	Yes
Country x trend	No	Yes	Yes	No	Yes	Yes
Province x trend	No	Yes	Yes	No	Yes	Yes
Province x country FE	No	Yes	Yes	No	Yes	Yes
Include irregular singles	Yes	Yes	No	Yes	Yes	No

Note: 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)-(3) and (4)-(6) 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 foreign spouse's country of origin. 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 specifications include country of origin, province, and year fixed effects; specifications in columns (2-3) and (5-6) include, in addition, linear trends interacted with province and country of origin fixed effects as well as province $\times$ country of origin fixed effects. In columns (3) and (6), the dependent variable is computed excluding (imputed) irregular immigrants at the denominator of the gains from marriage in equation (1). Standard errors clustered at the province level are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A4: Gains from marriage before and after the EU enlargements, difference-in-differences estimates allowing for announcement effects and for substitution across nationalities

	native male - foreign female		*native female - foreign male*	
	(1)	(2)	(3)	(4)
$EU2004 \times Post2003$	-0.011	-0.038	0.215	-0.071
	(0.073)	(0.081)	(0.188)	(0.204)
$EU2004 \times Post2004$	-0.224***	-0.244***	-0.099	-0.196
	(0.073)	(0.080)	(0.189)	(0.194)
$EU2004 \times Post2006$	-0.163**	-0.134*	-0380**	-0.519***
	(0.066)	(0.072)	(0.163)	(0.150)
$EU2004 \times Post2007$	-0.109*	-0.243***	-0.238	-0.191
	(0.064)	(0.075)	(0.149)	(0.192)
$EU2007 \times Post2003$	0.600***	0.543***	-0.671***	-0.697***
	(0.090)	(0.100)	(0.129)	(0.163)
$EU2007 \times Post2004$	0.494***	0.451***	0.151	0.192
	(0.082)	(0.093)	(0.135)	(0.147)
$EU2007 \times Post2006$	-0.402***	-0.502***	-0.020	-0.028
	(0.075)	(0.087)	(0.143)	(0.164)
$EU2007 \times Post2007$	-1.311***	-1.467***	-1.738***	-1.780***
	(0.075)	(0.090)	(0.134)	(0.205)
Observations	50,349	50,349	31,125	31,125
R-squared	0.612	0.781	0.691	0.852
Year, prov, country FE	Yes	Yes	Yes	Yes
Interacted FEs & trends	No	Yes	No	Yes

Note: This table shows the effect of admission to the EU on gains from marriage and cohabitation, 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)-(2) and (3)-(4) 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 foreign spouse's country of origin. The main explanatory variables are interactions between dummies for marriages between natives and new EU citizens ($EU2004$ and $EU2007$) and dummies for the periods after the announcement ($P2003$ and $P2006$) and implementation of each enlargement ($P2004$ and $P2007$). All specifications include year fixed effects, linear trends interacted with province and country of origin fixed effects, and province $\times$ country of origin fixed effects. Regressions 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 A5: Gains from marriage before and after the EU enlargements, difference-in-differences estimates (robustness to controlling for the number of singles OLS and IV)

	native male - foreign female (from country Y)				*native female - foreign male (from country X)*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$newEU \times PostEU$	-0.857***	-0.872***	-0.936***	-0.946***	-0.878***	-1.157***	-0.916***	-1.198***
	(0.066)	(0.069)	(0.086)	(0.090)	(0.105)	(0.166)	(0.108)	(0.164)
$\ln(Single\,Males)$	-0.991***	-1.243***	-1.000***	-1.253***	-0.808***	-0.783***	-0.842***	-0.978***
	(0.012)	(0.062)	(0.026)	(0.077)	(0.039)	(0.122)	(0.0410)	(0.149)
$\ln(Single\,Females)$	-0.749***	-0.672***	-0.750***	-0.609***	-1.151***	-1.632***	-1.189***	-1.618***
	(0.032)	(0.085)	(0.029)	(0.114)	(0.132)	(0.249)	(0.142)	(0.254)
$\ln(Single\,Males\,of\,country\,Y)$			0.0888***	0.0461				
			(0.027)	(0.122)				
$\ln(Single\,Females\,of\,country\,X)$							0.180***	0.542***
							(0.022)	(0.098)
Estimation	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Observations	50,349	33,739	44,721	30,751	31,125	17,837	29,539	17,252
R-squared	0.826	0.233	0.831	0.234	0.860	0.200	0.861	0.200

Note: This table shows the effect of admission to the EU, 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)-(2) and (3)-(4) present the results for couples formed by native males and females, respectively. The dependent variable is gains to marriage, computed as in equation (1), across cells defined by province, year, and foreign spouse's country of origin. 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 ($postEU$). We also include as explanatory variables the log number of single males and females. We instrument the number of single males and females by lagged immigrant population vectors (supplies) of males and females, respectively. OLS estimates are reported in odd-numbered columns, and IV estimates are reported in even-numbered columns. All specifications include country of origin, province, and year fixed effects, linear trends interacted with province and country of origin fixed effects as well as province × country of origin fixed effects. Regressions 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 A6: Gains from marriage before and after the EU enlargements, difference-in-differences estimates (robustness with local labor markets)

	native male - foreign female		*native female - foreign male*	
	(1)	(2)	(3)	(4)
$newEU \times PostEU$	-0.862***	-0.675***	-1.027***	-0.849**
	(0.0587)	(0.216)	(0.0941)	(0.104)
Observations	50,349	168,396	31,125	107,274
N. Marriages	3,402,324	3,268,522	3,239,135	3,156,421
Provinces	99		99	
Local Labor Markets		581		581
R-squared	0.809	0.614	0.845	0.757

Note: 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)-(2) and (3)-(4) 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 (or local labor markets), year, and foreign spouse's country of origin. 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 specifications include country of origin, province (or local labor markets), and year fixed effects, as well as linear trends interacted with province (or local labor markets) and country of origin fixed effects and province $\times$ country of origin fixed effects. Regressions are weighted by province (or local labor markets) population. Standard errors clustered at the province (or local labor markets) level are reported in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Internet diffusion and intermarriages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	native male, female from country y				native female, male from country x			
	log n. marriages		gains from marriage		log n. marriages		gains from marriage	
$\ln(SingleFemales_y)$	0.122***	0.152***						
	(0.015)	-0.018						
$\ln(SingleFemales_y) \times internet$		-0.002***						
		(0.000)						
$\ln(SingleMales_x)$					0.075***	0.092***		
					(0.022)	(0.017)		
$\ln(SingleMales_x) \times internet$						-0.001		
						(0.001)		
$newEUxPostEU$			-0.972***	-0.013			-1.072***	-1.988***
			(0.075)	(0.190)			(0.143)	(0.279)
$newEU \times PostEU \times internet$				-0.074**				0.069**
				(0.015)				(0.027)
Observations	39,313	39,313	39,313	39,313	23,793	23,793	23,793	23,793
R-squared	0.928	0.928	0.818	0.819	0.929	0.929	0.853	0.853
Interacted FEs & trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: 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 the 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 4. All specifications include year fixed effects, linear trends interacted with province and country of origin fixed effects, and province × country of origin fixed effects. Regressions 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 A8: Spouses' characteristics, before and after the EU enlargements

	(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

Note: This table shows the effect of the EU enlargements 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 enlargements. 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 A9: 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			

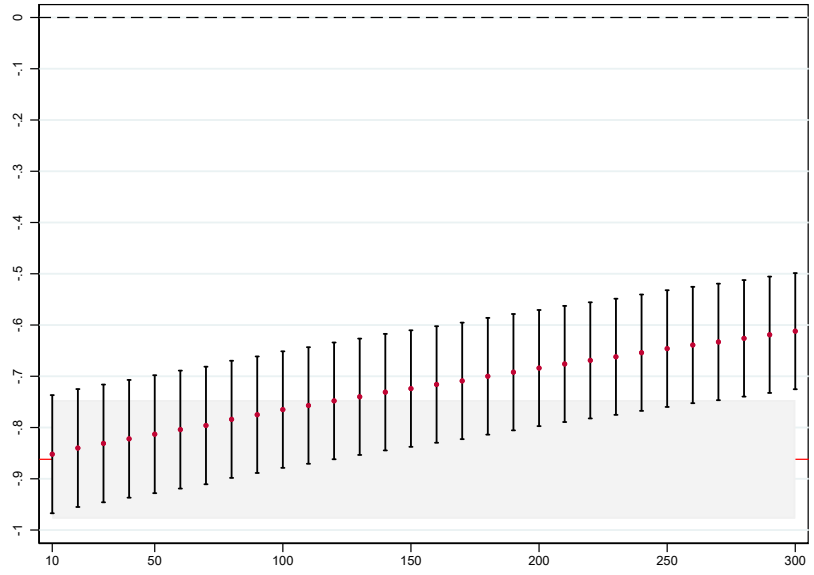
Table A10: ISMU survey, summary statistics for regular and irregular immigrants

	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			

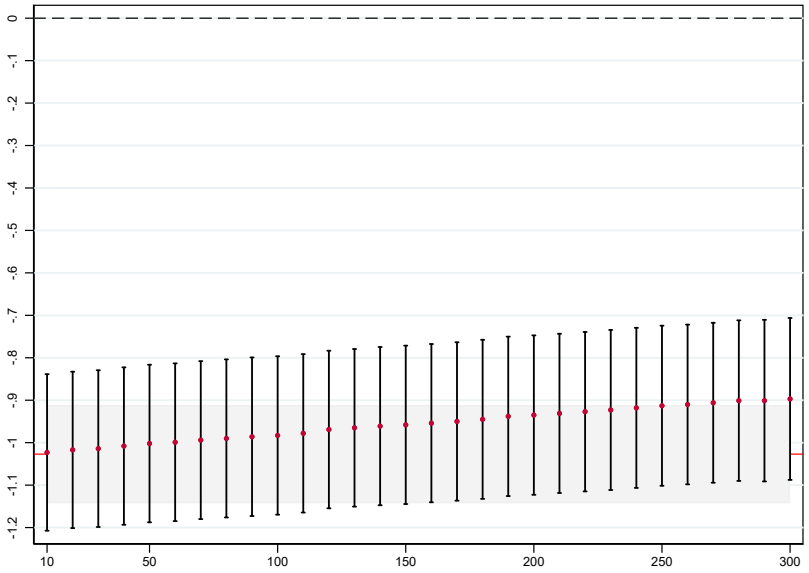
Note: ISMU (2001-2016) data, Lombardy. Individual level data.

Figure A1: Changes in gains from marriage between natives and new EU citizens – Robustness accounting for Measurement error in Unofficial Singles

(a) Native husband – foreign wife

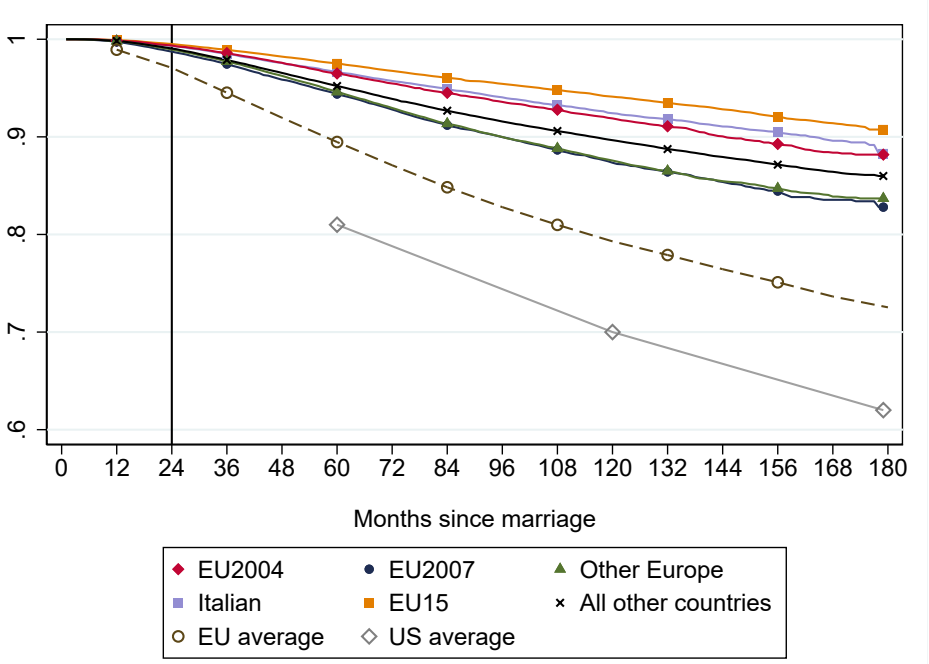


(b) Native wife – foreign husband



Notes: This figure shows the sensitivity of our estimated coefficient of interest (i.e., β in equation 2) and associated confidence intervals when progressively increasing the number of irregular immigrant singles in 10% steps over their (estimated) real number and allocating them at random across cells. The shaded area denotes the confidence interval of the baseline estimate.

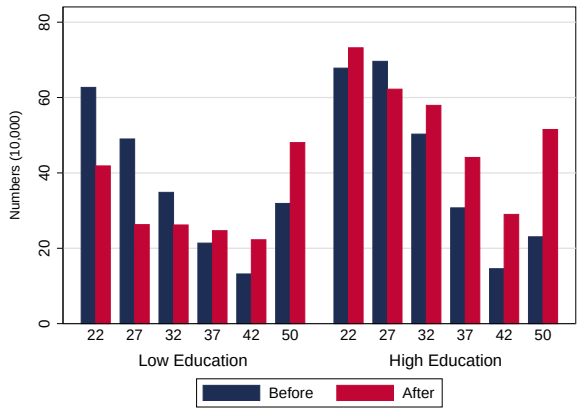
Figure A2: Survival function of marriages



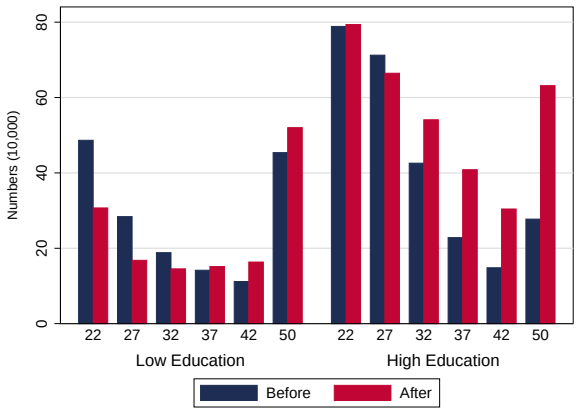
Notes: This graph shows the survival function of marriages in which (at least) one spouse was Italian, by area of origin of the other spouse. The vertical line indicates the period after which the foreign spouse can apply for Italian citizenship. In addition the graph reports the (average) survival function of marriages in Europe (EUROSTAT, 2012) and in the USA (National Health Statistics Reports, 2012).

Figure A3: Distribution of population vectors, by age and area of origin

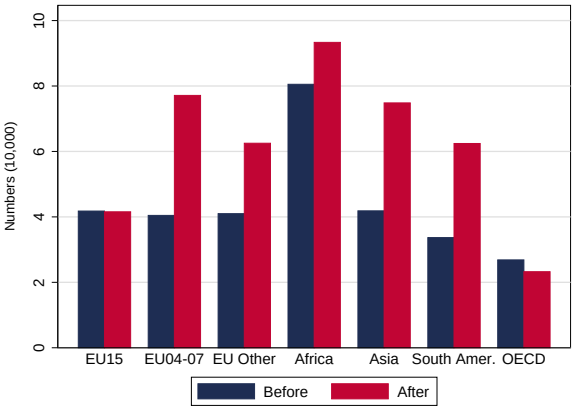
(a) Men by Age and Education



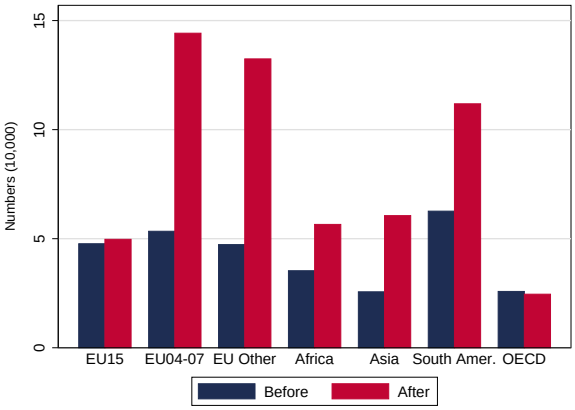
(b) Women by Age and Education



(c) Men by area of origin



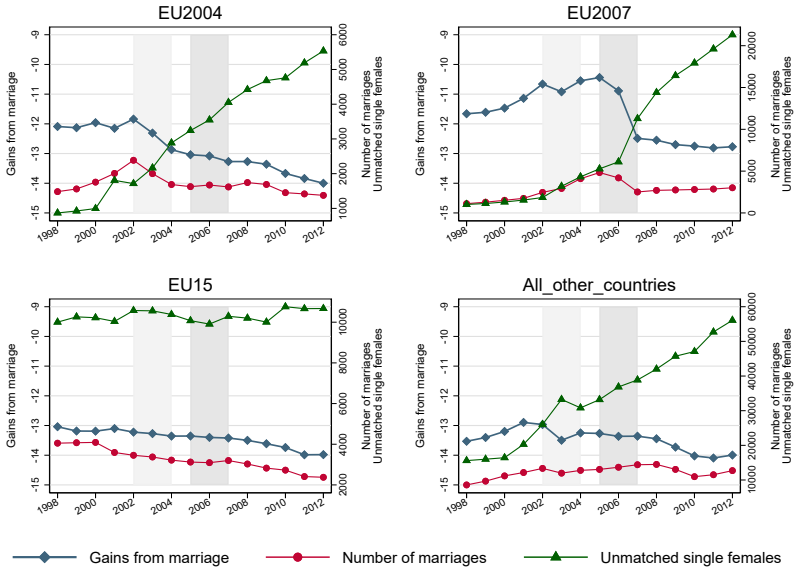
(d) Women by area of origin



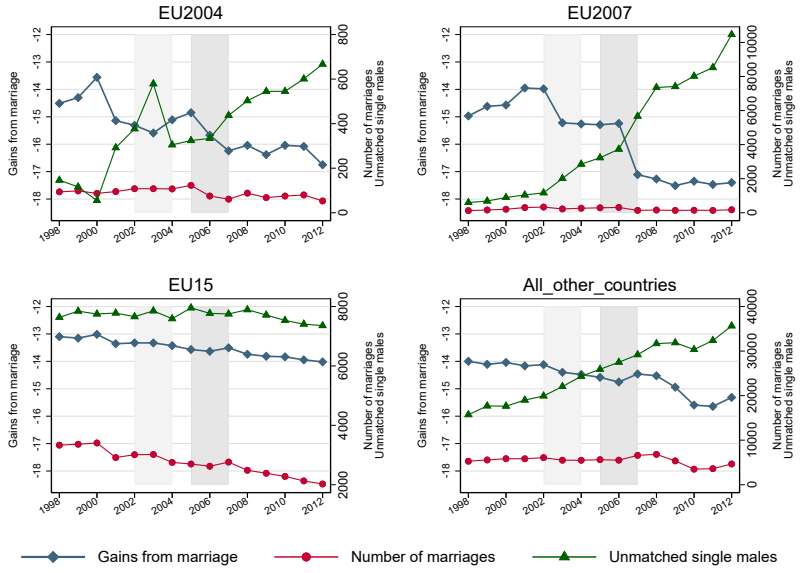
Notes: Marriage (1998-2002 and 2007-2012) and individual census (2001, 2011) data, Italy. The Figures report the distributions of men and women population vectors by age and education and area of origin, separately.

Figure A4: Singles, marriages, and gains from marriage by area of origin, 1998-2012

(a) Native husband – foreign wife

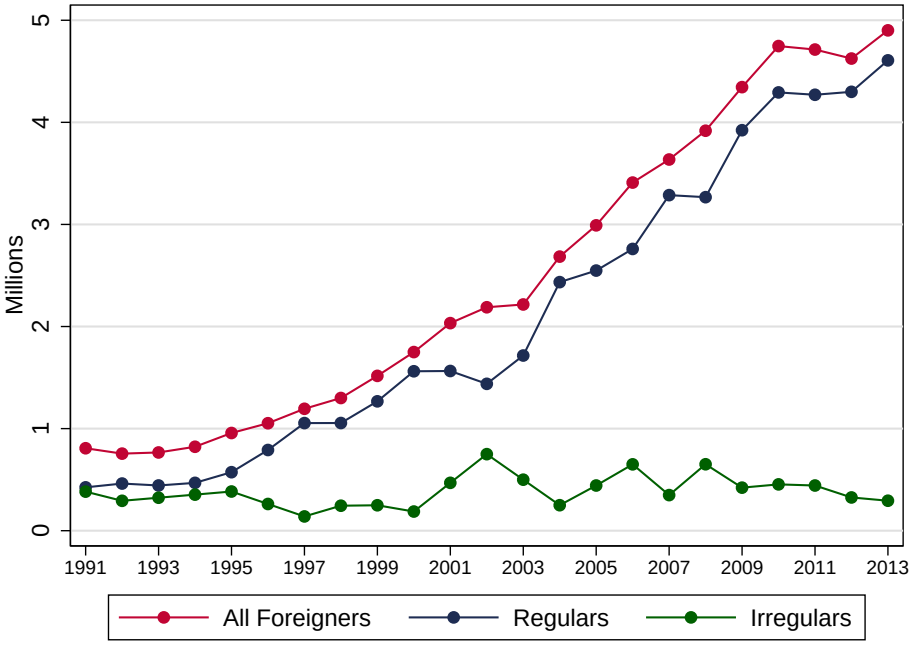


(b) Native wife – foreign husband



Notes: The graphs plot the number of 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 A1. *Source*: ISTAT, marriage records from vital statistics registries (1998-2012) and individual Census data.

Figure A5: Regular and irregular immigrants, 1991-2013



B Enlargement Process

The admission of a new country into the EU is a lengthy process, which may last for several years. The candidate member country may also be required to implement radical reforms to satisfy the Copenhagen criteria, which require transparent democratic institutions, protection of human rights and minorities, and a developed market economy. In addition, national laws must comply with the EU standards.

The first step in the admission process is the *Association Agreement* (AA), which allows the candidate country tariff-free access to some EU markets in exchange for political and economic reforms. If the AA is approved, the candidate country can submit its membership application, which the EU Council grants upon a favorable opinion. The ensuing negotiations concern more than 30 chapters, each one linked to a specific policy area (e.g., agriculture, education, competition). Finally, the accession treaty incorporates the resulting conditions (including transitional arrangements). The accession treaty must then be approved by the EU Parliament and the EU Council and ratified by the new member country through a parliament vote or a referendum.

As it should be expected, the whole process entails considerable uncertainty. Several candidate member countries – namely Serbia, Montenegro, Macedonia, Albania, and Turkey – have been negotiating admission to the EU for several years, and their cases are still far from settled. Regarding the EU2004 and EU2007 countries, negotiations started in March 1998 with six countries (Cyprus, Estonia, Hungary, Poland, the Czech Republic, and Slovenia), and in February 2000, with the six remaining countries (Bulgaria, Latvia, Lithuania, Malta, Romania, and Slovakia). The EU Council decision occurred on December 13, 2002, for the EU2004 countries, and on December 16, 2005, for the EU2007 countries, while the official admission dates were May 1, 2004, and January 1, 2007. Table B1 shows the timing of the admission process for EU2004 and EU2007 countries in detail.

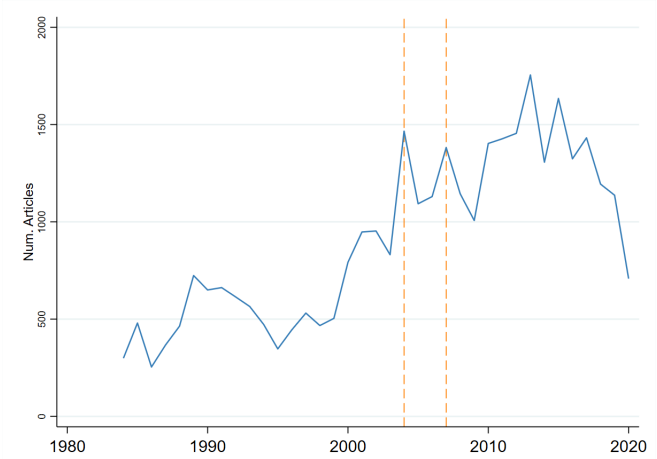
Table B1: Timing of admission process for EU2004 and EU2007 countries

Country	Euro Adoption	Schengen Area	Accession	Parliament	Referendum	Council	Negotiation	Application	EU AA in force	EU AA sign
Czech Republic	-	21 Dec 2007	1 May 2004	-	13 14 Jun 2003	13 Dec 2002	31 Mar 1998	17 Jan 1996	Feb 1995	Oct 1993
Malta	1 Jan 2008	21 Dec 2007	1 May 2004	-	8 Mar 2003	13 Dec 2002	15 Feb 2000	16 Jul 1990	Apr 1971	Dec 1970
Poland	-	21 Dec 2007	1 May 2004	-	7 8 Jun 2003	13 Dec 2002	31 Mar 1998	5 Apr 1994	Feb 1994	Dec 1991
Estonia	1 Jan 2011	21 Dec 2007	1 May 2004	-	14 Sep 2003	13 Dec 2002	31 Mar 1998	24 Nov 1995	Feb 1998	Jun 1995
Cyprus	1 Jan 2008	-	1 May 2004	14 Jul 2003	-	13 Dec 2002	31 Mar 1998	3 Jul 1990	Jun 1973	Dec 1972
Latvia	1 Jan 2014	21 Dec 2007	1 May 2004	-	20 Sep 2003	13 Dec 2002	15 Feb 2000	13 Oct 1995	Feb 1998	Jun 1995
Lithuania	1 Jan 2015	21 Dec 2007	1 May 2004	-	10 11 May 2003	13 Dec 2002	15 Feb 2000	8 Dec 1995	Feb 1998	Jun 1995
Hungary	-	21 Dec 2007	1 May 2004	-	12 Apr 2003	13 Dec 2002	31 Mar 1998	31 Mar 1994	Feb 1994	Dec 1991
Slovakia	1 Jan 2009	21 Dec 2007	1 May 2004	-	16 17 May 2003	13 Dec 2002	15 Feb 2000	27 Jun 1995	Feb 1995	Oct 1993
Slovenia	1 Jan 2007	21 Dec 2007	1 May 2004	-	23 Mar 2003	13 Dec 2002	31 Mar 1998	10 Jun 1996	Feb 1998	Jun 1996
Romania	-	-	1 Jan 2007	17 May 2005	-	16 Dec 2005	15 Feb 2000	22 June 1995	Jun 1995	Feb 1993
Bulgaria	-	-	1 Jan 2007	11 May 2005	-	16 Dec 2005	15 Feb 2000	18 Dec 1995	Dec 1995	Mar 1993

C Media coverage of the EU enlargements

Media coverage of the EU enlargements is both a proxy for attention towards this issue and the primary source of information available to immigrants about their legalization prospects. We scraped all articles from *La Repubblica*, one of the main Italian newspapers, from 1984 to (the first months of) 2021 – just less than a million articles – and selected all articles about immigration based on 145 keywords such as *Immigrazione, Immigrat*, ..., Extra-comunitar*, ... , Rifugiat**, and so on. We then count the number of articles per year mentioning at least one of the EU2004 or EU2007 countries concerning the EU enlargements, based on keywords such as *Allargamento, Unione Europea, UE*, and so on. Figure C1 clearly shows that attention towards these issues increased in the years close to the enlargement dates, and it spikes in 2004, 2007, and 2013 – the last corresponds to the admission of Croatia to the EU. These results are robust to using a different list of keywords and to removing one keyword at a time from our main list (results are available upon request).

Figure C1: Number of newspaper articles about EU Enlargements from 1984 to 2021



Notes: This figure shows the trend in newspaper articles about the EU enlargements' issues from 1984 to 2021. The sample refers to newspaper articles from *La Repubblica* related to immigration. To identify the EU enlargements, we exploit the following list of keywords = ['*adesione*', '*aderi*', '*anness*', '*entra*', '*far parte*', '*fara parte*', '*membro*', '*UE*', '*unione europea*', '*allargamento*', '*EU*']. The dashed lines refer to the years of the EU enlargements in 2004 and 2007.

D Imputation of the total number of singles

This section describes the procedure we use to impute the total number of singles, both among official and unofficial immigrants, at the denominator of the gains from marriage in equation (1). Throughout, we define singles as individuals aged between 18 and 60 who are not married.

D.1 Official immigrants

We obtained the exact number of singles by age, gender, nationality, and province from the Italian population Census in 1991, 2001, and 2011. We combined these data with migration inflows at the same level of aggregation during the period 1995-2012 to estimate the stock of immigrants by 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 immigrants 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 immigrants over time starting from $S_{c,2001}$ and adding up the yearly variation,

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

This procedure allows us to account for potential non-linear increases in the presence of singles over time. We implement the same procedure for the earlier period exploiting information on total immigrants and singles in census years 1991 and 2001 as well as migration inflows over the period 1995-2001.

D.2 Unofficial immigrants

As discussed in Section 3.2.2, neglecting unofficial immigrants 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) the 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 population Census; (ii) the number of singles and non-singles among irregular immigrants by cell in year 2002, from applications to the generalized amnesty in the same year; (iii) the 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 a method.

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:

$$\frac{S^I_t}{S_t} = \gamma \frac{M^I_t}{M_t}, \tag{12}$$

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:

$$\frac{S^I_{c,t}}{S^R_{c,t}} = \alpha_c \beta_t, \tag{13}$$

where α_c and β_t are cell-specific and time-specific constants.

In the next section, we provide indirect evidence consistent with these two assumptions. 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 (12):

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

Second, we aggregate individual-level data on regular and irregular immigrants by cell and

assume $\beta_{t^*} = 1$ to obtain

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

Third, we combine (12) and (13) 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

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

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 permit 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 the universe of applicants to the 2002 amnesty, which granted residence and work permits to about 700 thousand immigrants – the largest number ever legalized in Italy.[39] We use the total number of applications to estimate the 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.[40] 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 this probability using the survey data available from ISMU, already described in Section 2.3. As discussed there, the survey

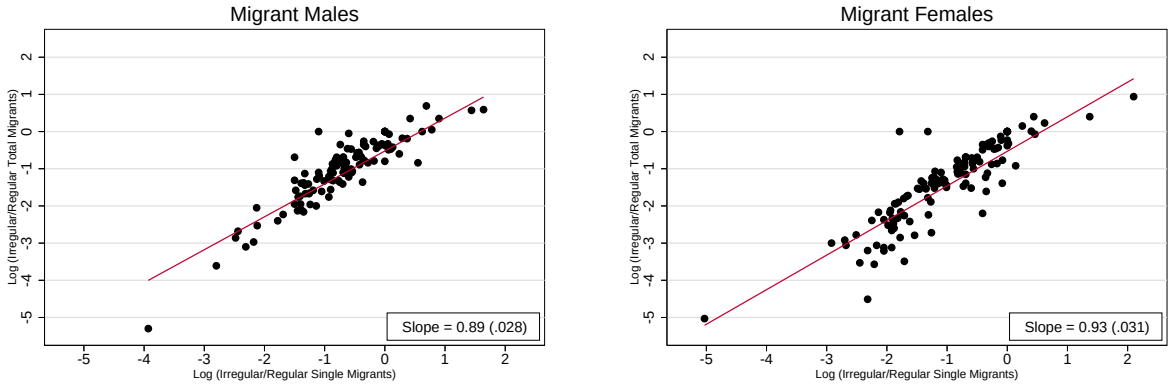
[39] We thank Francesco Fasani for kindly sharing these data; see Fasani (2009) for a detailed description of generalized amnesties in Italy.

[40] Table A9 provides some descriptive statistics, separately for male and female applicants – 377 and 318 thousand individuals, respectively.

provides detailed information on other individual characteristics, including gender, age, education, country of origin, marital status, employment, wage and province of residence. 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 A5), which we use to measure $M_t^I,\ \forall t \neq t^*$.

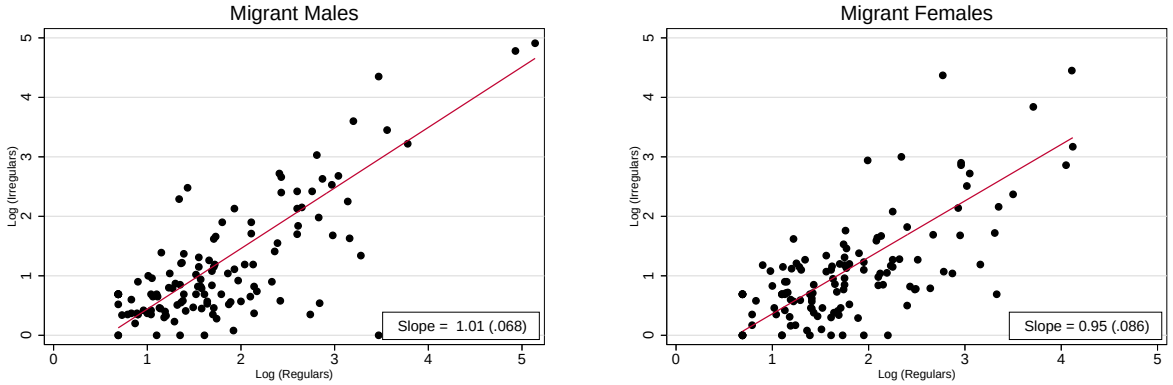
Indirect evidence. We present some empirical evidence to support the validity of our assumptions. On the one hand, we estimate the cross-sectional relationship between the ratio of irregulars over regulars among singles and total migrants across ethnic groups for the year 2002 of the amnesty. Figure D1 plots the cross-sectional log-log relationship for male and female migrants, separately. Figure D1 shows a strong positive and significant correlation between the ratio of single irregular migrants and the ratio of irregulars among total migrants, in line with our first assumption. On the other hand, we explore the log-log relationship between S^I and S^R exploiting variability across ethnic groups in year 2002. Figure D2 shows a strong positive association between the two variables of interest for both male and female migrants. Moreover, we do not reject the null hypothesis that the estimated elasticity is different from one, strongly supporting our second assumption of proportionality.

Figure D1: Test First Assumption: Regular and Irregular Migrants



Notes: This figure plots the relationship between the (log) ratio between the number of irregular over regular among *single* migrants against the (log) ratio between the number of irregular over regular among *total* migrants by ethnic group of origin, separately for male (left panel) and female (right panel)– after weighting by the number of provinces where each ethnic group is living.

Figure D2: Test Second Assumption: Regular and Irregular Migrants



Notes: This figure plots the relationship between the (log) number of single irregular migrants against the (log) number of single regular migrants by ethnic group of origin, separately for male (left panel) and female (right panel)– after weighting by the number of provinces where each ethnic group is living.

E Data and Sample Construction

We obtained restricted access to administrative Italian data at the individual level from ISTAT through its ADELE Laboratory.[41] In what follows, we start describing our data sources and variables of interest; then passing on to a discussion of the sample construction.

E.1 Marriage, Separation, Fertility

Marriage. We exploit marriage records from municipal vital statistics registries, which contain the universe of marriages celebrated in Italy from 1998 to 2012. Our data provide information about: i) the marriage, such as the date of marriage, the type of ceremony (religious or civil), the municipality of the ceremony and the choice of the property regime (community or separation property), and ii) the main socio-demographic characteristics of the spouses, such as the date and province of birth, the province of residence at the time of marriage, the previous marital status, the educational attainment, the employment and occupational status, and for immigrant individuals the nationality and the country of origin. We select marriages of adult individuals over 18 years old. In order to account for out-migration selection of families, the sample is restricted to marriages where at least one spouse is resident in Italy at the time of the marriage.

Our data cover the universe of marriages formed in Italy. Hence, we might potentially omit to consider those marriages that, despite being part of the Italian market, have been celebrated abroad. Of particular concern are homogamous marriages of foreign spouses, which for any reason, are more likely to be underestimated. The omission of foreign homogamous marriages might affect our estimates, in particular in case we systematically underestimate the number of homogamous marriages between new EU countries before the EU enlargements with respect to the after period. To have a sense of the importance of this bias, we recover the number of married couples living in Italy by exploiting marital status information from censuses 2001 and 2011 and we limit ourselves to marriages celebrated in the years right before the censuses. We compare the number of marriages in our registry data and the number of marriages in census data, and we correct our estimates by the difference in marriages observed between the two sources. Table E1 reports estimates of the missing

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

marriage rate.

Table E1: Missing Marriage Rates by Ethnic Group and Census Year

	Census 2001	Census 2011
EU15	0.86	0.53
EU2004/2007	1.52	1.36
EU_Other	1.78	1.55
Africa	1.74	0.97
Asia	1.63	1.22
South America	1.12	1.29
OECD Other	0.88	0.43

Separation. Separation data come from the registries of civil court chancelleries and cover the universe of legal separations registered in Italy in the period 1998-2012. Separation data provide information (among other things) on separation proceedings, date and place of marriage, and date and place of birth for both spouses. We focus on separation rates, which provide a more accurate representation of the timing of marital instability in the Italian context compared to divorces, for two main reasons. First, separation is the juridical act that starts the divorce proceedings. Second, with Law 74/1987 and until 2015, a minimum period of 3 years of legal separation was required in order to eventually obtain divorce.

Fertility. Fertility data come from municipality births registries, which contain the universe of individual birth records by municipality, for each year from 1998 to 2014. Individual birth records include socio-demographic variables of interest such as gender, date and province of birth, citizenship and parental information regarding their date of birth, citizenship and marital status.

E.2 Census Data on Singles and Cohabitations

Singles. We derive the distribution of singles from individual Census data of 1991, 2001 and 2011. We select adult individuals between 18 and 60 years old. We consider an individual as single in the case that she/he is never married, legally separated, divorced or widowed. Population Census data provide detailed individual-level information, such as gender, marital status, country of origin and nationality, age, educational level, and wealth.

Cohabitations. We exploit individual-level data from Census, which contain the universe of the population resident in Italy in 2001 and 2011. Individual Census data provide information about the family and the family id, the municipality of residence, and the main

socio-demographic characteristics including gender, age, marital status. For immigrant individuals, Census data provide additional information about the nationality, the country of origin, the year of arrival in Italy. We identify cohabitations, i.e., heterosexual love relations between partners who are living together but who are not legally married, by exploiting information on the family identifier, marital status and type of the relationship.

F Identification and Performance of the Estimator

We first show the ability of the estimation method to recover key model parameters. We then evaluate the performance of our estimator. We simplify the model for tractability and expositional clarity. We abstract from considering longitudinal variation, and we focus on a single cross-section of marriage data.

F.1 Identification in a Single Cross Section

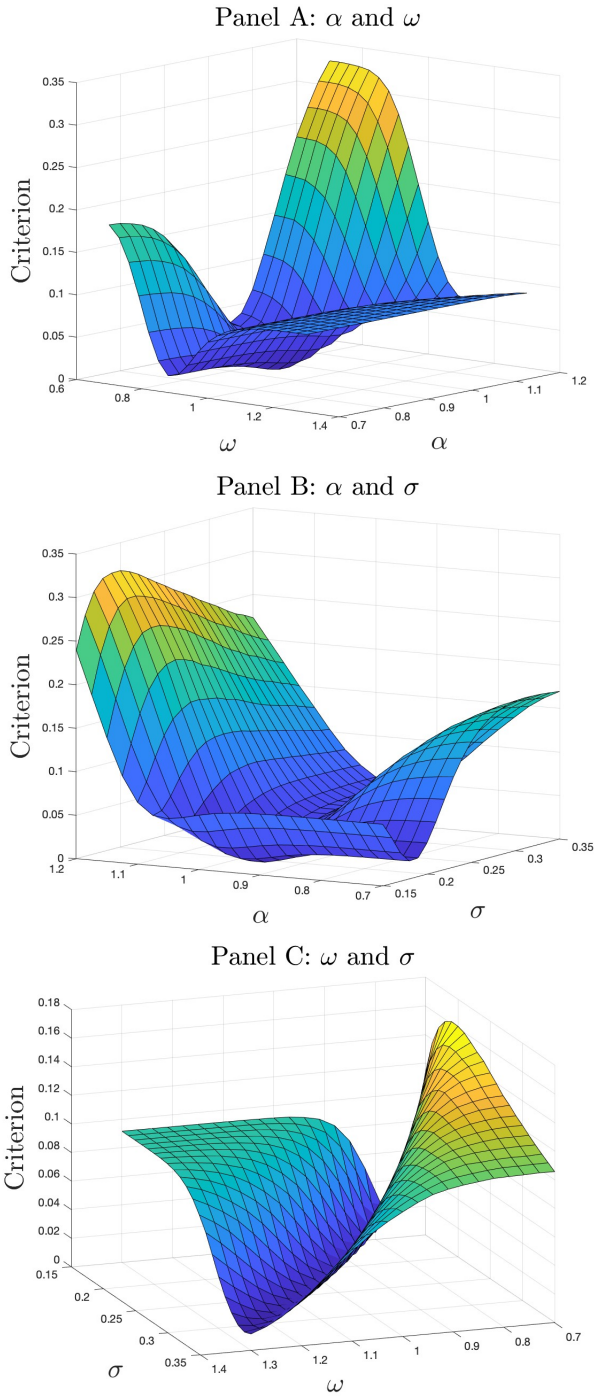
We consider only one market populated with I available men and J women. Individuals belong to two groups, that we label natives and foreigners. Individuals are subject to the same standard deviation of the "love" shock σ, and face the same outside value, ω. The marriage surplus associated with each type of marriage is denoted:

$$\phi = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}$$

The problem is then characterized by three parameters. The surplus is defined as $S_{ij} = e^{\varepsilon_{ij}} - 2\omega$ for homogamous marriages and as $S_{ij} = \alpha e^{\varepsilon_{ij}} - 2\omega$ for heterogamous ones. We define the "observed" data as the marriage patterns by nationality-pair groups, computed at a given value of the parameters, i.e. $\alpha^* = 0.95$, $\omega^* = 1$ and $\sigma^* = 0.25$ and for $I = J = 600$. These values are in the same order of magnitude than in the main estimation in the paper. The estimation criterion consists in the quadratic distance between the "observed" data and the prediction of the model for any value of the parameters on a grid. We show in Figure F1 the surface of this estimation criterion. The estimation criterium is unique and situated at the true value of the parameters. The estimation is able to clearly distinguish the systematic value of the surplus, α, from either the outside value ω or the standard deviation of the "love" shock in Panels A and B.

There is also only one set of values for σ and ω in Panel C that minimises the criterion function. However, there are a set of values for σ and ω that bring the criterium relatively close to zero situated in the valley in Panel C. Hence, in practice and in a single cross-section, it may be difficult to separately identify those two parameters. Further restrictions are needed to better distinguish their role in shaping marital choices. In the estimation performed in the main part of the paper, we leverage variation across marriage markets at the province level, and we restrict σ to be the same across all provinces, but we allow for province specific outside values. This helps to better pin down the "love" shock and the outside values.

Figure F1: Identification in a Single Cross Section



Notes: The figure plots the estimation criterion with only a cross-section of marriage data. The surplus for couple (i, j) is $S(i, j) = \phi(i, j)e^{\varepsilon_{ij}} - 2\omega$, with $Var(\varepsilon_{ij}) = \sigma^2$ and $\phi(i, j)$ equal to one or α for homogamous or heterogamous couples, respectively.

F.2 Monte Carlo Analysis

We now consider three different provinces $p \in \{A, B, C\}$. Within each province p, we assume that we observe N sub-markets populated with $I = J$ available men and women per sub-market. Individuals belong to two groups, that we label natives and foreigners. Within all markets, individuals are subject to the same standard deviation of the "love" shock σ, while the marriage surplus associated with each type of marriage possibly varies across provinces and it is denoted:

$$\Phi_p = \begin{bmatrix} 1 & \alpha_p \\ \tilde{\alpha}_p & 1 \end{bmatrix}$$

Provinces differ by a different outside value, denoted ω_p, with $p \in \{A, B, C\}$. In line with the discussion in Section F.1, having province specific outside values while having the same variance of the shock helps to pin down these two sets of parameters.

The estimation is done using the simulated method of moment (MSM), explained in Section 4.3 in the paper. The empirical moments consist of the marriage rates in each province p. To derive theoretical moments, we consider two different amounts of simulated individuals. In the first set of replications, we set the number of men and women $I = 300$ per sub-market and consider $N = 2$ sub-markets per province. This leads to a total of 1,800 simulated men and women that can potentially marry. In the second set, we use $I = 500$, $N = 10$ sub-markets per province, leading to 15,000 simulated individuals. To put these numbers into perspective, when estimating our model in the main text, we consider a market size of $I = J = 600$ individuals, and use 75,600 simulated men or women in total.

Let $\theta = [\omega_A, \omega_B, \omega_C, \sigma, \alpha_p, \tilde{\alpha}_p]$ denote the vector of parameters. The matching problem in its primal version is defined in equation (5). We solve the model for a given value of the parameters that we take as the true value of the parameters, (θ_0). This is done using $N = 20$ sub-markets per province p to eliminate simulation errors when determining the "true" marriage data. We then re-estimate the vector of parameters, θ, 250 times using a Nelder-Mead algorithm. We report the results in Table F1. The table reports the average departure from the true value of parameters and the standard deviation of the estimates; both expressed in percentages.

Table F1 shows that even with a small number of simulations, our estimation method is able to recover all the parameters of the model, see columns 2 and 3. Increasing the number of simulations brings the average of the estimates closer to θ_0 and reduces the dispersion of the estimates (columns 4 and 5). As the effect of legal status is identified by comparing surplus parameters across two periods in a given market, this precision translates as well to the effect of legal status in a panel setting.

Table F1: Monte Carlo Analysis

Parameter	True value θ_0 (1)	% deviation $E(\hat{\theta}-\theta_0)/\theta_0$ (2)	%. variation $Std(\hat{\theta})/\theta_0$ (3)	%. deviation $E(\hat{\theta}-\theta_0)/\theta_0$ (4)	%. variation $Std(\hat{\theta})/\theta_0$ (5)
α_A	0.9	-0.107 %	1.78 %	-0.209 %	1.29%
$\tilde{\alpha}_A$	0.97	0.0638 %	1.89 %	0.113 %	1.27%
α_B	0.97	-0.158 %	1.73 %	0.0558 %	1.2%
$\tilde{\alpha_B}$	0.9	0.519 %	1.75 %	-0.183 %	1.17%
α_C	0.85	1.15 %	1.71 %	-0.102 %	1.16%
$\tilde{\alpha_C}$	0.8	0.554 %	1.59 %	0.64 %	1.24%
ω_A	0.85	-0.0892 %	0.787 %	0.149 %	0.582%
ω_B	0.9	0.0657 %	0.783 %	0.292 %	0.42%
ω_C	0.95	0.556 %	0.706 %	0.321 %	0.394%
σ	0.25	0.444 %	0.793 %	0.472 %	0.447%
Individuals I per sub-market		300		500	
Number of sub-markets N		2		10	
Total Men & Women $I \times N$ in 3 prov		1,800		15,000	

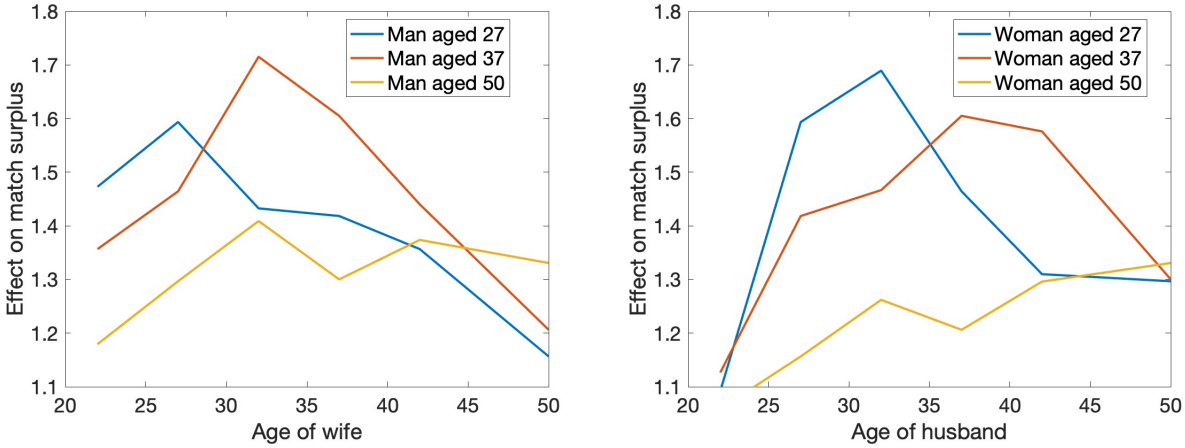
G Additional Estimated Parameters

G.1 Additional Estimated Parameters

We start by presenting the parameters related to age, which are displayed in Figure G1. 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 G1 displays the parameters about education and home ownership. We find mild assortative mating along educational lines, 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 wealth, as proxied by home ownership. For example, when both partners are Italian, having one 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 similar effects, or slightly higher magnitudes, if one of the partners is non-Italian. The probability of owning property among immigrants is relatively low, especially among young adults. For immigrants 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-thirds. The results in Table G1 have two implications. First, given the asymmetry in the surplus associated with housing, mixed couples are more likely to include a native

Figure G1: Estimated parameters, marriage surplus by age and gender



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

Table G1: Estimated parameters, marriage surplus linked to education and home ownership

		Women					
		Education					
		Low	High				
Men	Low	1.1	0.883	Housing, natives			
	(0.0001)	(0.0001)					
	High	0.939	1	No	Yes		
			(0)	-			
			No	1	1.77	Housing, non natives	
				-	(0.0006)		
			Yes	1.6	1.77	No	Yes
				(0.00011)	(0.00301		
					No	1	1.11
						-	(0.00397)
					Yes	1.19	2.04
						(0.00099)	(0.00428)

Note: The table displays the part of the surplus associated with education and housing, as defined in equation (6). Asymptotic standard errors are displayed in parentheses.

Table G2: Parameters of the fertility equation

	Coeff.	Sdt err
γ_0	-0.3640	(0.0147)
γ_1	0.0323	(0.0014)

Note: The table show the parameters and standard errors determining fertility; see equation (8).

man and a foreign woman. Second, given the rate of ownership by groups, this surplus component favors marriages between older native men (with home ownership) and younger foreign women (without home ownership). Finally, the increase in the surplus compensates for part of the cultural and age differences among those couples.

Table G2 displays the parameters determining fertility (see equation (8)). The parameter γ_1 is positive but small. This implies that couples with a higher marital surplus are more likely to have children within a window of three years, but that other couple characteristics also drive fertility.

G.2 Marital surplus and cultural distance measures

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. Interestingly, we relate our revealed measure to other measures of cultural distance 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 (WVS). Finally, genetic distance provides an indication of the degree of genealogical relatedness between two populations (Cavalli-Sforza and Piazza, 1994; Spolaore and Wacziarg, 2009, 2016). It is defined based on the coancestry coefficients: the heterozygosity index, i.e., the probability that two alleles from a given locus, selected at random in two populations, will be different. Thus, the greater the genetic distance between two populations, the longer they have been apart from each other, and the greater would be the difference in cultural values. We refer to Spolaore and Wacziarg (2016) for an accurate description of these four measures of cultural distance.

Figure G2 displays how our measure of the cultural affinity of the marital surplus relates to these alternative measures. As measured in these four different ways, greater distance along cultural lines is systematically associated with a lower residual surplus from marriage. However, our surplus measure is only weakly related to linguistic and genetic differences. Those two measures only explain 5 to 9 percent of the variation in the surplus. Instead, cultural proximity based on the WVS and religious distance explain each 15 percent of the variance. For those two measures, a one standard deviation increase in cultural distance decrease the marital surplus by about 0.4 standard deviations.

Finally, we exploit longitudinal variation over repeated waves of the WVS to examine whether the EU enlargements brought a convergence in different cultural values (listed in Table G3) between citizens of new and older EU countries, respectively. However, this does not seem to be the case. Figure G3 plots distance in the average values observed in EU15 countries and three groups of countries, namely EU2004, EU2007, and other European countries, in 1999 and 2010 (i.e., before and after the EU enlargements, respectively). EU2004 and EU2007 countries do not seem to converge to EU15 values, neither in absolute terms nor relative to other European countries. We formally test for convergence in values by estimating the following equation:

$$DV_{it} = \sum_t \delta_t (newEU_i \times postEU_{it}) + FEs + \varepsilon_{ijt}, \tag{14}$$

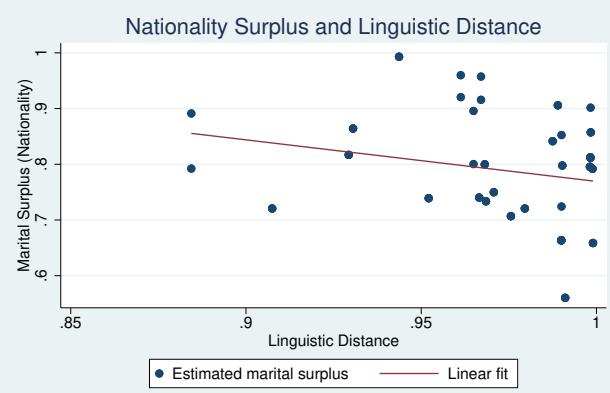
where DV_{it} is the distance in values between country i and EU15 countries in year t; the δ_t's are wave-specific coefficients capturing differential convergence of new EU countries to EU15 cultural values over time, compared to other European countries; and the other terms

on the right-hand side are defined as in the main estimating equation (2). Figure G4 shows no differential convergence in cultural values between new EU and EU15 countries, neither before nor after the enlargements.

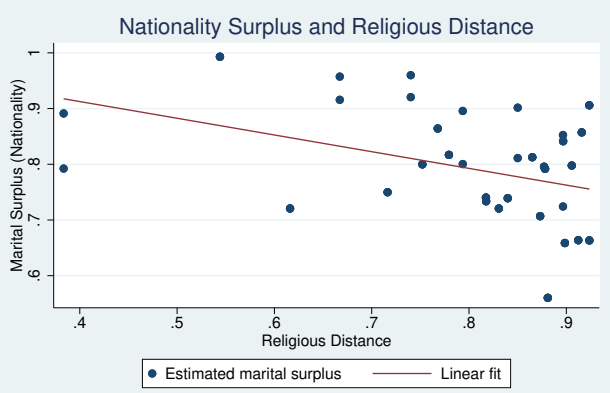
Based on these findings, we conclude that the EU enlargements did not bring convergence in cultural values, at least in the short term, which implies, in turn, that changes in outcomes around the enlargements can be attributed to the effect of legal status acquisition as opposed to cultural convergence.

Figure G2: Marriage surplus and cultural distance

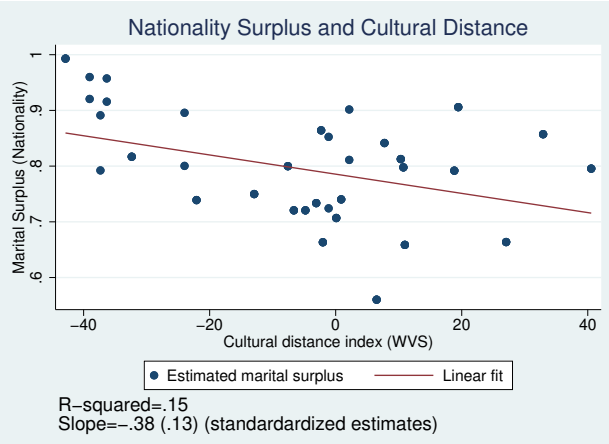
(a) Linguistic Distance



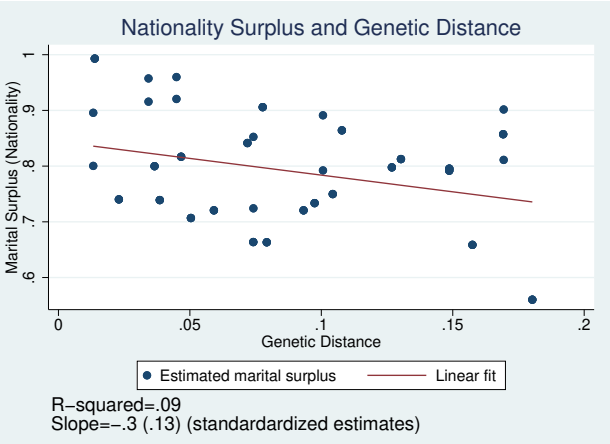
(b) Religious Distance



(c) WVS Distance



(d) Genetic Distance



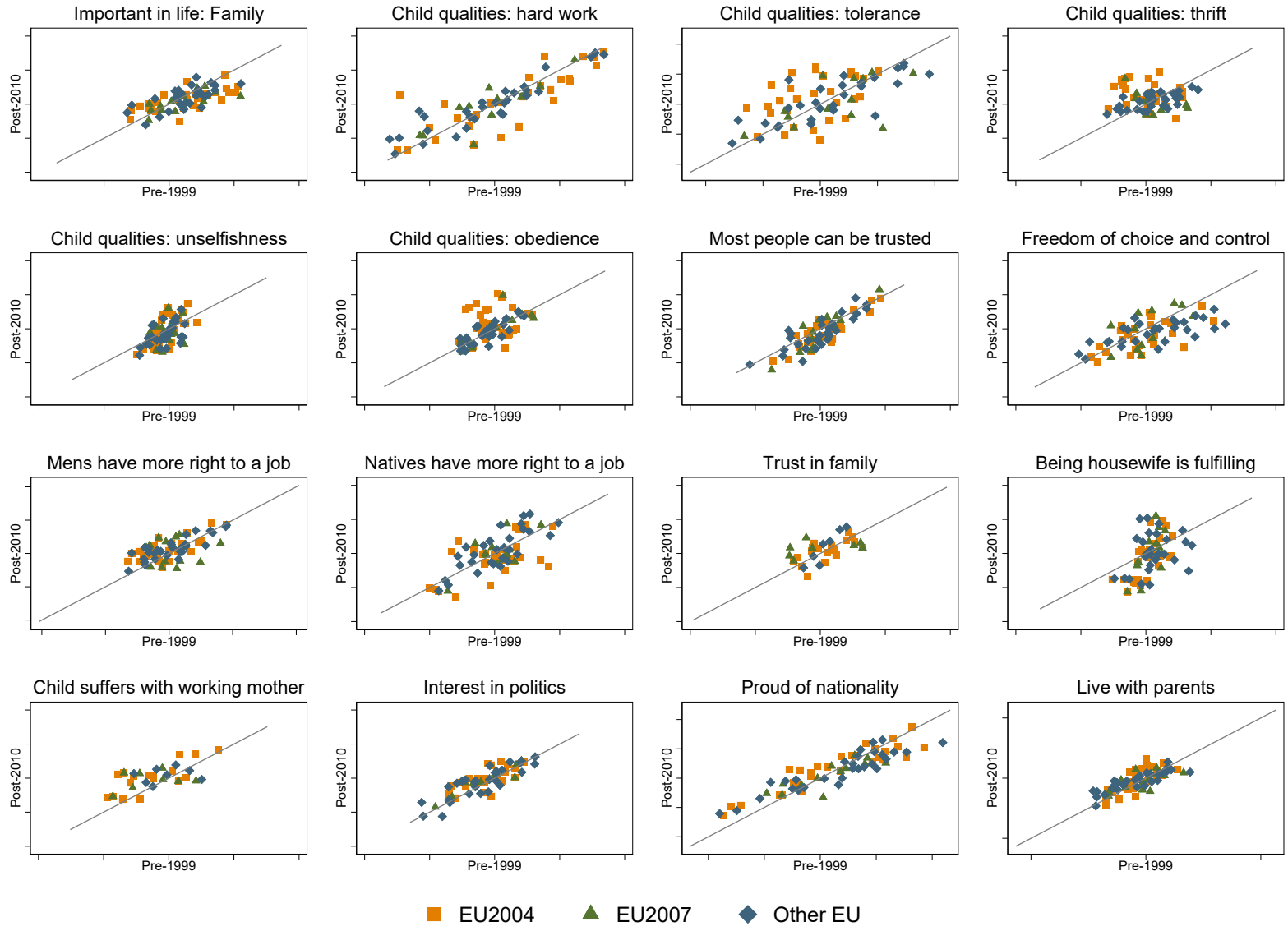
Notes: The Figure plots the relationship between four ex-ante measures of cultural distance from Spolaore and Wacziarg (2016) and the estimated cultural-based component of marriage surplus estimated from the model; see equation 7 and panel C of Table 9.

Table G3: World Value Survey List of Variables

Variable	Question	References in the Literature
Perceptions of Life		
A001	Important in life: Family	Alesina and Giuliano (2010, 2011)
A030	Important child qualities: hard work	Tabellini (2010)
A035	Important child qualities: tolerance and respect for other people	Tabellini (2010)
A038	Important child qualities: thrift saving money and things	Guiso et al. (2003);Tabellini (2010)
A041	Important child qualities: unselfishness	Tabellini (2010)
A042	Important child qualities: obedience	Alesina et al. (2015); Tabellini (2010)
A165	Most people can be trusted	Tabellini (2010);Alesina and Giuliano (2011); Guiso et al. (2003)
A173	How much freedom of choice and control	Tabellini (2010)
Work		
C001	Jobs scarce: Men should have more right to a job than women	Alesina and Giuliano (2010);Alesina et al. (2013); Guiso et al. (2003)
C002	Jobs scarce: Employers should give priority to (nation) people than immigrants	
Family		
D001	How much do you trust your family (4 categories)	Alesina and Giuliano (2011)
D057	Being a housewife just as fulfilling	Alesina and Giuliano (2010); Guiso et al. (2003)
D061	Pre-school child suffers with working mother	
Politics and Society, and National Identity		
E023	Interest in politics	Alesina and Giuliano (2011)
G006	How proud of nationality	Figlio et al. (2019);Hofstede et al. (2010)
Demographics		
X026	Do you live with your parents	Alesina et al. (2015)

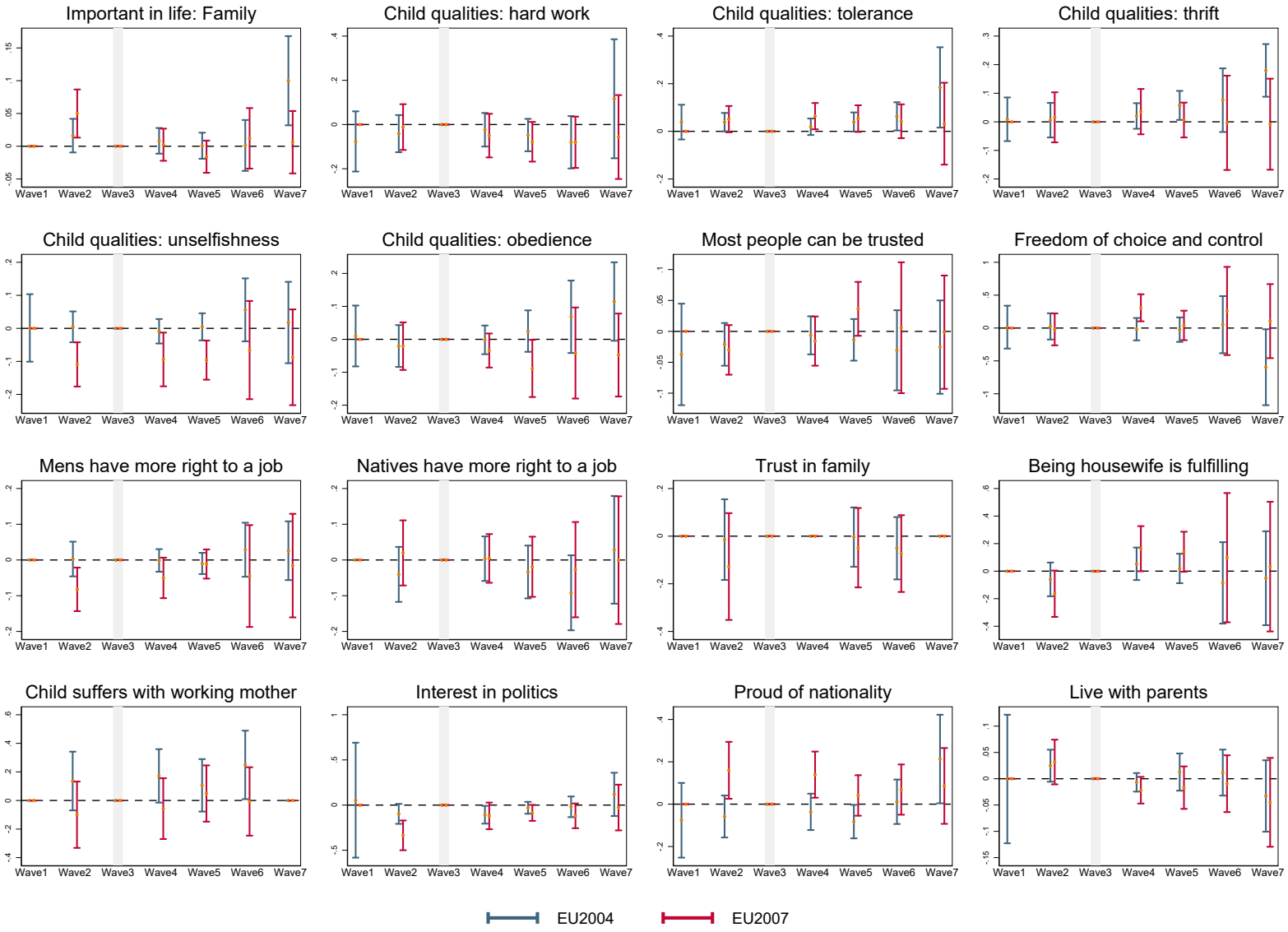
Note: This table presents the list of variables and questions we focused on in our analysis about values and attitudes on various topics, such as perceptions and important aspects of life, work, family and gender roles, traditions, trust, and attitudes towards other groups. The table highlights some notable references in the literature. Source: Integrated Worlds Value Survey (WVS) and European Value Survey (EVS) dataset (1981-2020).

Figure G3: Average cultural distance of European countries before and after the EU enlargements



Notes: This figure plots the correlation of mean country-pair culture distances before (pre-1999) and after (post-2010) the EU enlargements for various measures describing cultural values and attitudes. The complete list of cultural variables of interest is reported in Table G3. The analysis focuses only on European countries, including EU15, EU2004, EU2007, and EU Other countries. The 45-degree line is reported in black. Source: Integrated Worlds Value Survey (WVS) and European Value Survey (EVS) dataset.

Figure G4: Average cultural distance before and after the EU enlargements, difference-in-differences estimates



Notes: This figure shows the estimated effect of admission on cultural distance by country-pair over time. The main explanatory variables are interactions between dummies for EU2004 and EU2007 countries and a full set of wave fixed effects. The graphs plot the estimated coefficients and associated confidence intervals, based on standard errors clustered at the country-pair level. The shaded area corresponds to the reference category (wave 3, 2005-2009).