Location Choice, Labor Market Conditions, and Marital Sorting among Immigrants

Evidence from Germany

Jacek Barszczewski*

December 22, 2022

Click **here** for most recent version.

Abstract

Labor market integration programs affect not only immigrants' economic outcomes but also their social integration. In this paper, I analyze how measures of social integration, like the share of marriages with natives and immigrants' spatial concentration, change under different policy scenarios. I first show correlations between immigrants' labor market outcomes, marital patterns, and spatial distribution. Then, using German data, I estimate a structural model with location, marriage, and labor supply decisions. The model reflects two trade-offs immigrants face: a) partner choice: "marry your like" vs. economic gains from marriage with a native, and b) location choice: a region with higher wages vs. a region with better marriage opportunities. Model simulations reveal that: 1) reducing the immigrant-native income gap by 25% decreases immigrants' spatial concentration (by 2.9%), but lowers the share of immigrant women married to natives (by 2 pp); 2) declining the regional wage gap by 50% significantly reduces immigrants' spatial concentration (by 15%), increases the share of immigrant men married to native (by 1.1 pp), but decreases the share of immigrant women married to natives (by 0.6 pp). I also find that ignoring adjustments in location and marriage choices under both policies overstates the decrease in immigrant-native income inequality and underpredicts the welfare gains. The reason for that is when immigrants' labor market position improves, they give up part of their income gains and marry natives less often to satisfy their taste for similarity in partners' origin, increasing their welfare.

Keywords: Immigration, Integration, Marriage, Location Choice, Labor Market
JEL codes: D10, J12, J31, J61, R23

*PhD Candidate, Universitat Autònoma de Barcelona and Barcelona School of Economics. Email: jacek.barszczewski@bse.eu. Website: www.jbarszczewski.pl. This research has received funding from the European Research Council (ERC) under the European Union's Horizon2020, research and innovation program GA project number 804989-DYMOLAMO. I thank my advisor Joan Llull, as well as Nezih Gunner, Limor Golan, George-Levi Gayle, Hanna Wang, Ada Ferrer-i-Carbonell, Serdar Birinci, Robert A. Pollak, Oksana Leukhina, Hannah Rubinton, Victoria Gregory, Paulina Restrepo-Echavarria, SangMok Lee, Prasanthi Ramakrishnan, Karol Madoń, Katherina Thomas and seminar participants in Barcelona and St. Louis.

1 Introduction

Over the last decade, more than 45 million people permanently migrated to OECD countries[2], increasing the share of the foreign-born population by 16%[3]. This growing number puts a spotlight on immigrant integration in the public debate. In response, policymakers implement various programs that support integration. The majority of these programs[4] focus on integrating immigrants into the labor market. If successful, they improve the immigrant economic situation. Despite the economic dimension of integration, there is also the social one. The intensity of immigrant interactions with natives also contributes to their level of integration. Among others, the social dimension is measured by the frequency of marriages with natives and immigrant spatial concentration (Lazear, 1999, Danzer and Yaman, 2013, Boeri et al., 2015, Cutler et al., 2008). What are the consequences of labor market integration policies for the social dimensions of integration? How do the immigrants' marital patterns change? Do immigrants adjust their location choices? To what extent do those changes impact income inequality and welfare?

The answer to these questions is not trivial and depends on several factors. While searching for a partner, immigrants face the decision to marry another immigrant or a native. On the one hand, in most OECD countries, foreign-born earn less, on average, than natives. In this sense, intermarriage[5] may improve their financial situation. On the other hand, immigrants, likewise natives, show preferences for similarity, which makes other immigrants more attractive. Consequently, in one scenario, improving immigrants' labor market outcomes raises their attractiveness in native eyes, increases the intermarriage rate, and fosters immigrant social integration. In the alternative scenario, it decreases relative gain from cohabitation with a native, leads to more marriage between immigrants, and mitigates the positive effect of labor market policy. Moreover, marital patterns depend on the partners' availability, which, in turn, depends on location choices. Immigrants can trade regions with higher wages for regions with more immigrant partners, assuming they have a taste for similarity. As a result, depending on the character of changes in the labor market and preferences for similarity, immigrants might adjust location choices in a way that leads to a decrease or increase in their spatial concentration. Further, the direction of changes in location and marriage patterns impact household resources and thus affect income inequality and welfare.

---

[2] Author's calculation for years 2010-2019 based on *International Migration Outlook 2021*

[3] Ibid.

[4] The most common integration programs are active labor market integration policies, i.e., language training, labor market training and work practice, subsidized employment, and job search assistance.

[5] The existing literature does not uniquely define the *intermarriage* term. In general, *intermarriage* refers to marriage outside own social group. It has traditionally been restricted only to actual formal marriage. Nowadays, this way of defining intermarriages seems to omit the other common forms of partnership. Possibly due to social pressure, immigrant-native couples even more often avoid a formal framework (Benson, 1981). It stresses the need to extend the *intermarriage* definition to other forms of partnership. The other problem emerges with the definition of own social group. (see Rodríguez-García (2015) and Elwert (2018) for further discussion). This paper uses the term *intermarriage* as an informal and formal partnership of foreign-born and native-born individuals. By *nonintermarriage*, I denote any other form of an informal and formal

To analyze potential consequences of labor market integration policies and capture the relevant trade-offs, I build a structural model with an equilibrium marriage market in which immigrants and natives choose their location, find partners, and optimize their labor supply. I estimate the model with German microdata and quantify the effect of labor market policy outcomes on marriage and location patterns. Further, I conduct welfare and income inequality analyses to understand how controlling for adjustments in marriage and region choices changes the initial economic effect of the labor market policies. I propose the modeling approach that allows for answering the research questions and conducting relevant analyses in three ways. First, location choices depend on labor market conditions so that I can predict the spatial concentration of immigrants under different labor market policy scenarios. Second, while choosing a partner, agents take into account future household income. Due to that, I can simulate how different labor market conditions change marriage choices in equilibrium. And third, in my policy exercises, I can carefully control for interdependence between location and marriage choices.

The model presented in this paper builds on the recent works on the matching models by Chiappori et al. (2017), Adda et al. (2020), Galichon and Salanié (2021) in the spirit of Becker (1973, 1974). Agents make a labor supply decision within the static collective household framework. Natives and immigrants differ in wages and leisure preferences by education (college vs. noncollege) and region (North, South and West)[6] to capture the observed variation in income and labor supply. I allow wages to vary by marriage status and by partner's origin. By that, I account for potential immigrants' wage premium from intermarriage empirically shown by, i.e., Meng and Gregory (2005), Basu (2015), Elwert and Tegunimataka (2016). Further, the marriage surplus depends on the future household budget, which generates differences in marital gains by the partner's education and origin. Finally, agents have preferences toward similarity in origin and education to capture the observed assortative mating in marriage patterns.

Natives and immigrants make lifetime location choices based on regional characteristics, such as expectations towards marriage and labor market outcomes and the value of local amenities. Since locations differ in the level of wages by origin, education, and gender, agents have incentives to distribute disproportionally across regions. As a result, the underrepresented types have more bargaining power in the local marriage market and benefit from the higher transfers in the matching process (I model the marriage market in a frictionless framework with transferable utility). The transfer sizes impact the expected utility of settling in a region and influence its attractiveness in equilibrium, directly linking location and marriage choices. This link is especially crucial for immigrants since their number is relatively small, so any changes in spatial distribution have a profound impact on marriage outcomes (van Tubergen and Maas, 2007, Harris and Ono, 2005, Choi and Tienda, 2017). Region's utility also includes the exogenous amenity index, which I create, following Dia-

partnership.

[6]Following the Federal Statistical Office of Germany, I define four macro-regions: South - Hesse, Baden-Wuerttemberg, and Bavaria; West - North-Rhine-Westfalia, Rhineland-Palatinate, Saarland; North - Schleswig-Holstein, Hamburg, Lower-Saxony, Bremen and Berlin. East region is dropped from analysis due to a very small migration population.

mond (2016) in the estimation process, based on the broad set of variables, i.e., access to public transport or number of severe crimes.

Thanks to the model structure, I conduct the estimation in three steps, starting with the household problem. It is a standard static labor supply problem, so labor market and leisure parameters are identified by observed variations in wages and labor supply choices. I fit this part of the model to the data from the German Socioeconomic Panel, waves 1984-2018. In the second step, I use consistent estimates of the labor market parameters to predict the total household economic gain. Thanks to that, I can later estimate tastes for similarity and endogenous transfers on the marriage market for the baseline scenario. The marriage market equilibrium conditions entirely determine the intrahousehold allocation of the economic gains for all possible matches. I identify the partner preference parameters by observed marital outcomes, following the approach by Choo and Siow (2006). In the last step, I use marriage parameter estimates to predict the expected utility of participating in regional marriage and labor markets. This value, together with the amenity index, determines the location choices. Next, I identify the taste for amenities by across-cohort variation in location choice probabilities. For the last two steps, I fit the model using the German Microcensus 2006, 2010, and 2015.

I find a significant gap between the earnings of immigrants and natives by gender and education. This finding is in line with the less than perfect international transferability of human capital (i.e., Chiswick and Miller (2009)). As a result, households with immigrants have lower disposable income than those with only natives of the same education level. Moreover, the estimated immigrant-native wage gap varies across regions. It means that, to some extent, the difference in the distributions of immigrants and natives across space is driven by variations in labor market outcomes. Next, the expected economic surplus generated by each type of household depends on household income and the value of leisure. Keeping the same level of partners' education, I find that the surplus is higher in the case of immigrant-native households due to higher preferences for leisure among immigrants (complementary effect). It makes mixed unions more attractive from an economic point of view.

Marriage market equilibrium conditions, preference parameters, and expected economic surplus from marriage determine agents' marital choices. Estimated similarity parameters imply that agents prefer to match with partners of the same origin and education. However, the preferences for similarity are stronger in origin than in education. I also find that estimated endogenous transfers between agents show some patterns in the bargaining power of agents in the marriage market. On average, agents with higher potential earnings and are more scarce in the population have a better negotiating position. As a result, they obtain higher endogenous transfer in the marriage market. It means that agents have incentives to choose a location with a lower wage but fewer agents of the same type and compensate for the loss in income by higher marriage market transfer. In this way, marriage market conditions partially counteract labor market motives for location choices.

Subsequently, I use the estimated model parameters to quantify the effect of labor market integration policies on intermarriage and spatial concentration of immigrants. I do so by simulating two counterfactual scenarios. In the first scenario, a government introduces a country-wide policy that reduces an immigrant-native wage gap (i.e. publicly available language courses). As a result, the increase in immigrants' wages equals 25% of the initial value of the gap by region, gender, and education. This increase is equivalent to an average wage rise of 4.6% for foreign-born men and 7.5% for foreign-born women. Under the second scenario, I assume that a government targets the regional variation in immigrants' wages (i.e. locally subsidize employment). The policy increases the earnings of non-native residents in regions with an overall lower income level. In the aftermath, the differences between the region with the highest wages and the remaining ones reduce by 50% of the initial value of the gap by gender and education level. This pay rise is equal to a 3.5% increase in the average wages of immigrants.

First, I show that outcomes of introduced policies lead to a decrease in the spatial concentration of immigrants[7]. The decline is more substantial in the case of a reduction in regional variation in wages and is equal to around 10% for noncollege- and 32% for college-educated immigrants. Therefore, to a different extent, both scenarios ease financial incentives to concentrate in the region with the most favorable labor market. It means that there is a positive impact of analyzed policies on social integration via adjustments in location choices. Next, I find that the effects of both policies on marriage patterns are mixed and vary by gender. Immigrant men are less likely to stay single (from -1.6 to -8.2 pp, stronger effect when the immigrant-native gap is reduced). Further, they are also more likely to be intermarried (from 0.9 to 2.2 pp). In the case of women, the increase in the number of marriages is smaller, and the probability of marrying native men decreases (up to -2.4 pp). As a result, the outcomes of analyzed labor market policies have a positive impact via intermarriages on the social integration of men but a negative (to a greater extent when the immigrant-native gap is reduced) in the case of women.

Finally, I conduct welfare and income inequality analyses. The policies that increase immigrants' wages mechanically reduce income inequality between immigrants and natives[8]. However, I find that ignoring the adjustments in marriage and location choices leads to overprediction of a decline in income inequality. In the case of reducing the immigrant-native gap and regional wage variation, the decline equals, respectively, 5% and 6%. Unlike income inequality, I show that ignoring both adjustments is associated with the underprediction of welfare gains. While reducing the immigrant-native gap, the underprediction equals 12%. In the case of the reduction in regional wage variation, it is even higher and equals 15%. It means that when immigrants' earnings rise, they give up part of their marriage economic gains by marrying natives less often. They do so to satisfy their taste for similarity in their partner's origin, increasing their welfare gains.

This paper is related to several strands of literature. Most closely related are studies on the integration of immigrants. By conducting immigrant-native income inequality and welfare analyses, I extend the literature that studies the effects of government integration programs on the economic performance of immigrants, see among others Hayfron (2001),

[7]I measure the spatial concentration by the total variation distance between uniform and observed distributions

[8]I measure income inequality between immigrants and natives as a percentage difference in the per capita income by gender and education.

Lochmann et al. (2019), Joona and Nekby (2012), Cohen-Goldner and Eckstein (2008, 2010). Hayfron (2001) and Lochmann et al. (2019) study the participation of immigrants in language, while Joona and Nekby (2012) evaluate whether intensive counseling and coaching improve immigrants' employment opportunities. Cohen-Goldner and Eckstein (2008), and Cohen-Goldner and Eckstein (2010) finds a positive effect of local training on wages and labor market participation. I also quantify the effect of labor market integration policies on intermarriage and spatial concentration of immigrants. It allows me to evaluate if those policies positively impact not only the economic integration of immigrants but also its social aspects. By that, I contribute to the literature that focuses on the determinants and socio-economic consequences of non-labor aspects of integration, see among others Kalmijn and van Tubergen (2006), Dribe and Lundh (2011), Chiswick and Houseworth (2011), Grossbard and Vernon (2020), Xie and Gough (2011), Min Zhou and Logan (1989).

My analysis of the marriage market builds on previous equilibrium models with a transferable utility, such as Chiappori et al. (2018) and Calvo et al. (2021). My model structure is closest to the one proposed in Chiappori et al. (2018) regarding the marriage market and household behavior. However, I focus on immigrant integration. Hence, I distinguish individuals not only by education level but also by immigrant status. Further, the first choice in my model is location decision instead of the decisions of human capital (education) investments. The paper by Calvo et al. (2021) relates to mine in that they focus on the relationship between labor and marriage markets and estimate their model using the same German data. In my model, a labor market impacts marriage patterns through changes in economic gains. Unlike, they examine how the connection between labor and the marriage market affects home production and patterns of job matching.

Combining the location choice decision with the marriage and labor markets is a novel feature of my model. On the one hand, my location choice decision model is inspired by the tradition of spatial equilibrium models initiated by Rosen (1979) and Roback (1982) and recently popularized by Diamond (2016). Unlike all these models, the equilibrium clearing in my model occurs in the marriage market. However, given the nature of my counterfactual exercises, I abstract from labor market equilibrium for tractability reasons. Unlike in these models, however, the availability of potential partners of different types plays a crucial role in my model. On the other hand, equilibrium marriage models generally focus on a single global market. However, several reduced-form papers show that marriage market outcomes differ across space and impact location choices, see among others Costa and Kahn (2000), Compton and Pollak (2007), Chiswick and Houseworth (2011). I allow for the endogenous spatial allocation of individuals on all sides of the market, which leads to changes in bargaining power, which fundamentally affect marriage market outcomes. The paper that similarly uses a setting with endogenous sorting and marriage market is Fan and Zou (2021). Contrary to them, I distinguish individuals by origin, so I can study the effects separately for immigrants and natives instead of focusing on the determinants of the spatial distribution of economic activities.

Finally, a few papers analyze the marriage patterns of immigrants and natives in an equilibrium framework. The most notable example is Adda et al. (2020). In that paper, the authors explore the trade-off between mating along cultural lines and legal status acquisition, which can positively impact labor outcomes. Adda and coauthors also study local marriage markets, but they take a geographical distribution of immigrants as given. Using my framework, I can investigate how the spatial concentration of immigrants would change given the anticipated changes in the labor market and what consequences it has for marriage patterns. The mechanisms associated with location choice are even more important while analyzing immigrants since: (a) they are more mobile than natives, so they might stronger respond to changes in location conditions; (b) they are a relatively smaller group compared to natives, so any change in the local composition of the marriage market has a more substantial impact on their marriage outcome.

The rest of the paper is organized as follows. Section 2 contains basic statistics and empirical facts linking location choice, marriage market, and economic integration of immigrants. Section 3 presents the model, while Section 4 discusses the data used and employed estimation strategy. Section 5 contains outcomes of conducted counterfactual scenarios. Section 6 concludes.

2 Descriptive evidence

In this section, I provide evidence linking location choice, marriage, and labor market outcomes. I use this evidence to motivate the research question and the model structure presented in the next section. To conduct the empirical analyses, I use German data. I choose to focus on Germany as it is an attractive country for immigrants from different origins. Immigrants from East European and Post-Soviet countries are the biggest immigration group, and their share in total migration stock is slightly above 30%. The second biggest group is Turkish immigrants, which share is equal to 17%. A similar share of immigrants is of Balkan origins. The last significant group is Southern European immigrants, which share is equal to 11%. The remaining 25% of immigrants came from other countries. The diversity of immigrants' origins creates a suitable environment for analysis that allows answering the research questions of this paper.

Figure 1 compares region choices and the difference in mean wages between immigrants and natives. To obtain the proper comparison, I use the difference in immigrant and native probabilities of settling down in one of three German regions. By that, I measure the relative overrepresentation of immigrants in the local population. The intuition behind this exercise is as follows: immigrants from different groups settle down more often compered to natives in regions where their wages are relatively higher. The figure suggests a positive correlation. The fitted regression indicates that closing the wage gap between immigrants and natives in a particular group by 1 pp (percentage point) leads to a 0.196 (s.e. 0.090) pp increase in differences in the probability of region choice in this group. In summary, labor market conditions could be an essential factor driving immigrants' location choices.

Beyond the difference in labor market conditions, regions also differ in the local social structure. Those differences may play a vital role in determining marriage patterns. Figure 2 presents the correlation between the share of intermarried immigrants and two characteristics

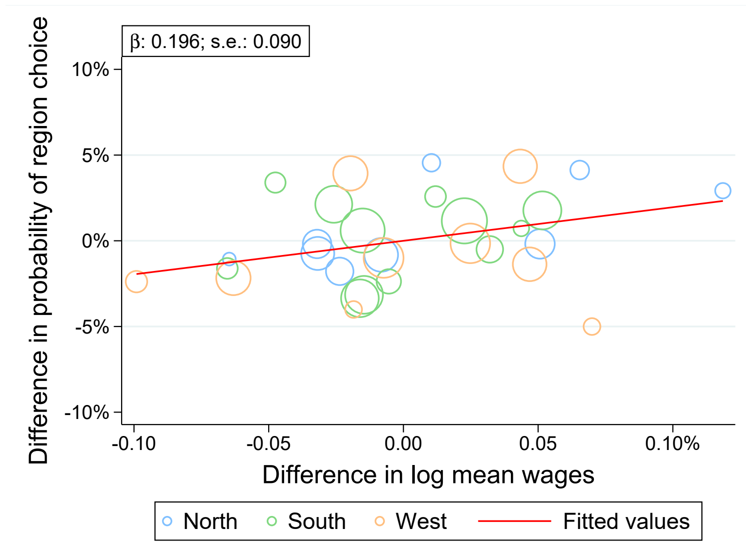


Figure 1: Difference in mean wage and spatial sorting

Notes: Each circle on the graph represents one group characterized by birth cohort, gender, education, and region of residence. The size of the circle corresponds to the size of the observation unit. Both variables are net of gender, education, cohort, year, age, and region fixed effects. Observations with a mean income difference above the 90th percentile and below the 10th percentile are dropped. *Source*: GSOEP 1984-2017 & German Microcensus 2006, 2010, 2015.

of the local marriage market: sex ratio in the immigration population (panel A); share of immigrants within the local opposite-sex population (panel B). The figure suggests a positive correlation between the intermarriage rate and the sex ratio. 1% increase in the sex ratio leads to a 0.858 p.p. (s.e. 0.399) increase in the intermarriage rate among males. The outcome indicates that a higher sex ratio leads to tougher competition in the marriage market for male immigrants. As a result, it provides incentives to search for a partner outside their origin group. This conclusion does not apply to female immigrants. In their case, the correlation is negative but insignificant. It means, that competition seems to play a more important role only for male immigrants. The downer panels of the Figure 2 suggest that the share of immigrants in the different sex local population negatively correlates with the probability of intermarriage. Suppose a share of females (males) increases by one p.p. In that case, the percentage of intermarried male immigrants decreases on average by 0.449 p.p. (0.573 p.p) with 0.089 (0.105) s.e. It implies that the bigger pool of immigrant partners lowers the probability that an immigrant finds a partner among natives. As a result, immigrants could consider those differences between regions while deciding on their future living place.

Reports on immigrants (i.e., OECD (2020)) suggest that they differ from natives regarding labor market outcomes. Those differences manifest later in disparities in the disposable income of households. Figure 3 presents differences in mean income between two types of households: mixed households, cohabitation of an immigrant and native, and all-immigrant households, where both partners are foreign-born. I conduct the analysis from the point of view of an immigrant and separately by gender and education level. The figure suggests

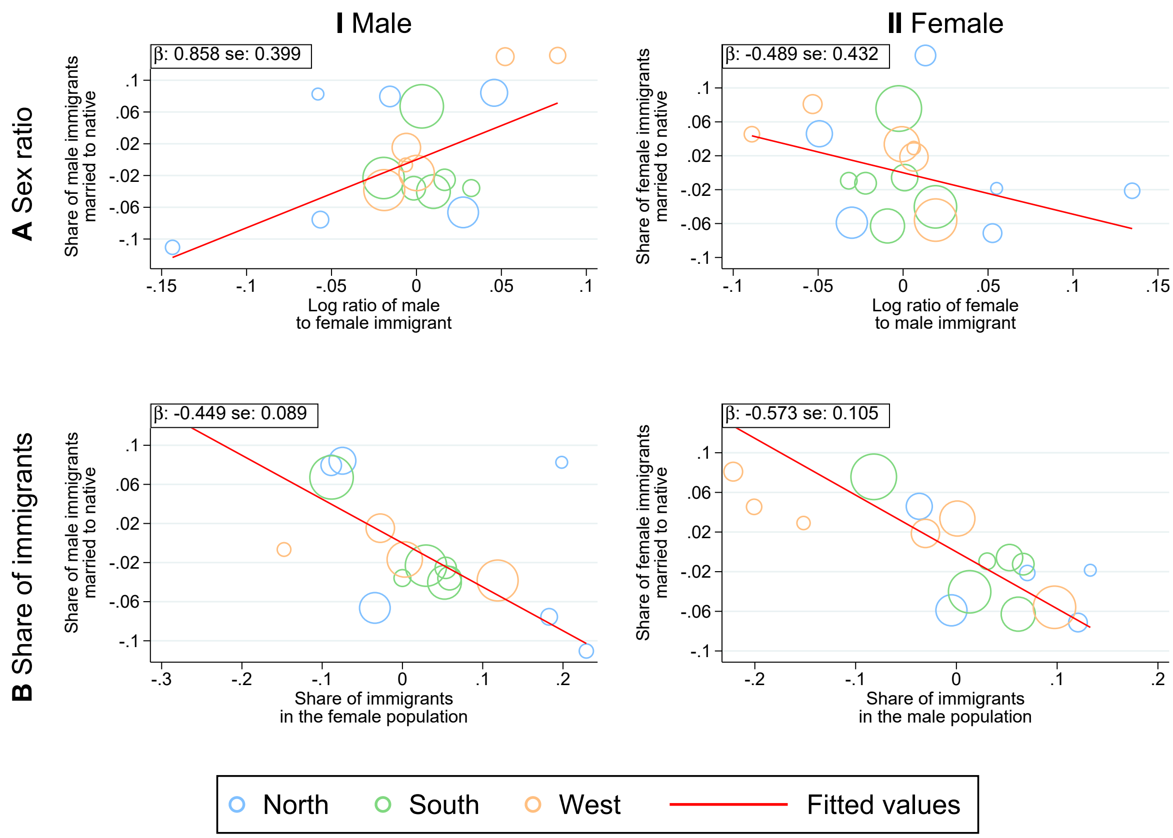


Figure 2: The correlation of intermarriage rate with the share of immigrants and sex ratio

Notes: Each circle on the graph represents one group. The groups are defined by birth cohort, gender, education, and region of residence. The size of the circle corresponds to the size of the group. I calculate the intermarriage rate as a share of immigrants married to natives in the group. The sex ratio is the number of males (females) per female (male). Intermarriage rate, sex ratio, and share of immigrants are net of the cohort, education, and region fixed effects. The lines represent the fitted regression lines, which slopes and their standard errors are included in the upper-left corner of each subplot. *Source*: Microcensus 2006, 2010 and 2015.

that, on average, mixed households are characterized by higher income than all-immigrant households. It links the marriage decision with economic well-being.

The correlations presented in Figure 1, 2, 3 are suggestive of the link between location decision, choice of partner, and labor market outcomes. Impact on the latter might influence the first two and change the final effect of the integration policy. As a result, endogenizing location and marriage choices in the equilibrium framework shed new light on the unintended effect of the pro-integration labor market policy.

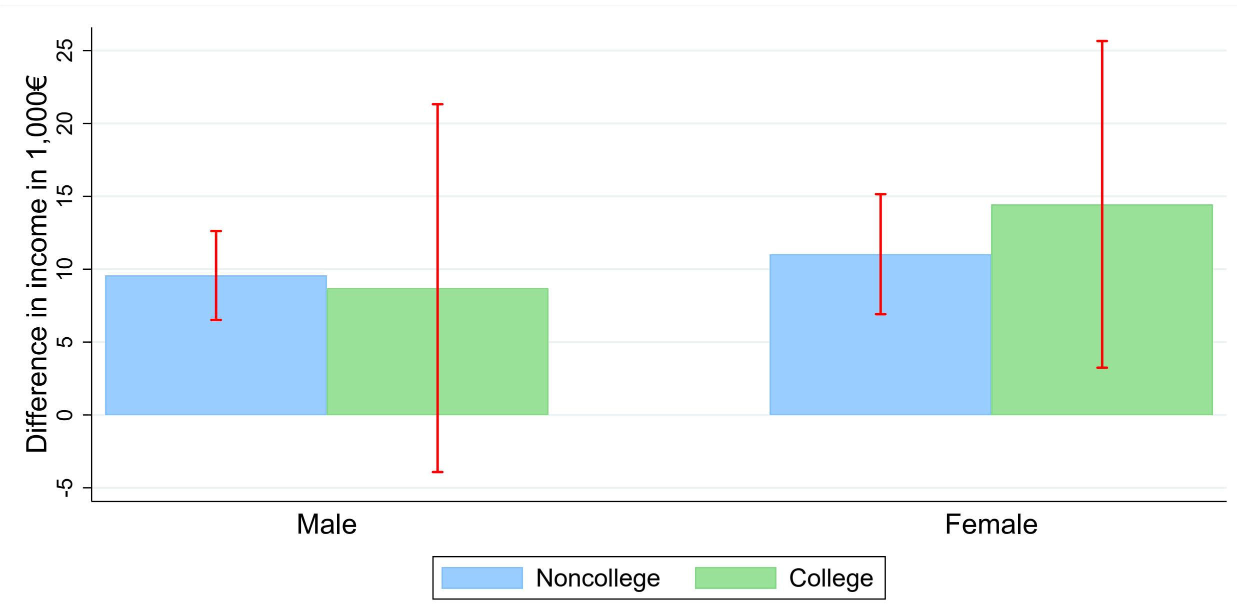


Figure 3: Difference in mean income between mixed and all-immigrant households

Notes: Each bar presents the difference in means of yearly household income expressed in thousands of EUR between households where only one member is native and households where both members are immigrants. The immigrant position defines gender and education. Household income is a sum of labor and nonlabor incomes. Both incomes are net of age profile, region, and year fixed effects. The red lines represent 95% confidential intervals for the calculated mean. The standard error of the mean is calculated using clusters at the household level. *Source*: GSOEP 1984-2017.

3 Model

Why do labor market conditions influence immigrants' marriage choices? The answer to this question can be briefly described. People marry for both economic and noneconomic reasons. Regarding pecuniary motives, couples can collect more resources than single agents. The size of a family's income depends on, among others, the partners' origins. Natives tend to earn, on average, more than immigrants, so households with them have higher disposable income. Further, married agents tend to perform better in the labor market, which is associated with the marriage premium described in the literature. The premium can differ by partner's origin. The size of the additional premium received by immigrants thanks to marriage with natives might depend on the labor market integration.

Regarding noneconomic reasons for marriage, people have a taste for similarity. The taste may play an essential role in immigrants' case since they can have preferences to marry somebody who shares similar values, language, or religion. As a result, immigrants can trade economic perspectives for cultural similarity. However, the possibility of trade-off depends on the local marriage market structure. The fewer immigrants in a different sex marriageable group, the harder to marry another immigrant and comparatively easier to marry a native. The structure of the marriage market is not exogenous but depends on immigrants' location choices. While choosing where to live, immigrants take into account two aspects. First, the economic situation in the region, in particular, the level of wages. Second, the number of desirable potential partners. Changes in the labor market situation in regions can lead to a stronger or weaker spatial concentration of immigrants, directly impacting marriage market conditions.

Four things are necessary to capture the abovementioned mechanisms: (1) the model of marriage and household behavior, (2) heterogeneity in origin among agents, (3) an endogenous location decision, and (4) wages varying by agent's origin, marital status, and spouse's origin. This list motivates the following setup.

3.1 Set-up of the model

In the model, agents belong to a cohort of women $\mathcal{F}$ or men $\mathcal{M}$. Each agent's life is divided into three stages, indexed 1-3. At the beginning of stage 1, agent of gender $g \in \mathcal{F} \cup \mathcal{M}$ posses a human capital H. It comprises two elements: origin and education. I denote agents' origin by $o \in \mathcal{O} \equiv \{n, i\}$, where n stands for native and i stands for immigrant. Agent is also characterized by education level denoted by $e \in \mathcal{E} \equiv \{e^1, e^2\}$. As a result, human capital can be expressed as a two-element set $H \equiv \{o, e\}$. The distribution of human capital has finite support $\mathcal{H}$ of cardinality 2×2.

At stage 1, all agents first draw a vector of location preferences. Then, they make lifetime decisions regarding a region of residence. Agent chooses location, denoted by r selecting from the set $\mathcal{R} \equiv \{r^0, r^1, r^2\}$. Region choice depends on local amenities and future marriage and economic perspectives. As a result, at the end of stage 1 agent lives in the region r, where next enters a marriage market to search for a partner.

At stage 2, agents draw a vector of marital preferences and then participate in the local marriage market chosen at stage 1. The agents match based on the level of human capital (education and origin) and marital preferences in the frictionless framework. An individual can marry a person of different sex, with origin $o*$ and education $e*$. Partner's human capital is then consistently denoted by H^*. The couple can be of 16 marriage types (four types of men and four of women). I denote the married couple's type by (H, H^*), where H is the human capital of the husband and H^* is the wife's. A single household's type of man and women with human capital H is denoted by $(H, \varnothing)$ and $(\varnothing, H)$, respectively. Marriage is a lifetime decision, so the outcome of stage 2 remains forever - there is no possibility of divorce or separation.

At stage 3, agents realize their productivity and leisure shocks and observe their wages and leisure preferences. Then, all households choose the optimal consumption of private and public goods and labor supply. I assume married couples make a Pareto efficient decision.

Agent's utility splits into three parts corresponding to the model's three stages. The first part comprises the working-life utility at stage 3, derived from the consumption of goods and leisure. The second part is the utility derived from participating in the marriage market. Finally, agents derive utility from regional amenities. My description of the model is as follows. First, I define the household maximization problem at stage 3. Then I describe the marriage market, taking working-life utility as given. Finally, I provide a brief description of the location choice.

3.2 Working-life utility of agents

At stage 3, agents choose the optimal consumption and labor supply levels. The choice is made based on observed wages and nonlabor income. Then, agent of gender g and human capital $H = \{o, e\}$ married to agent of origin o^* in region r earn wage given by:

$$w = W_g(H, o^*, r) \cdot \varepsilon = \exp\{\theta_{0g}(H) + \theta_{1g}(H, o^*) + \theta_2(H, r)\} \cdot \varepsilon \tag{1}$$

where:

$$\varepsilon | g, e \sim i.i.d.\ \mathcal{N}\left(0, \sigma^2_{\varepsilon|g,e}\right) \tag{2}$$

The exponential expression in Equation 1 represents the deterministic part of agents' wages. It consists of three components. The first component $\theta_{0g}(H)$ corresponds to the agent's human capital market value. To capture the gender wage gap in labor income, I let human capital market value differ for men and women. Equation 1 also allows immigrants and natives with the same education to have different human capital market values for two reasons. First, immigrants can have a different intercept than natives, which captures the effect of country-specific skills, like language (Llull, 2018). Second, natives and immigrants can differ in return to education (Borjas, 1985). The second component of Equations 1 $\theta_{1g}(H, o^*)$ represents market value shift relate to partner origin o^*. By that, the model allows for heterogeneity in agents' wages by spouse's origin. The wage premium associated with a partner's origin varies by gender to reflect empirical facts in the literature Meng and Gregory (2005), Meng and Meurs (2009). If agent is single, so $o^* = \varnothing$, then $\theta_{1g}(H, \varnothing)$ can be interpreted as a shift in market value due to being unmarried. The third component $\theta_2(H, r)$ is introduced to capture regional variation in earnings. It makes some regions more attractive due to better labor market conditions. Agents' wages are subject to independent and idiosyncratic productivity shock ε, conditionally on gender and education, normally distributed across agents with zero mean and variance $\sigma^2_{\varepsilon|g,e}$.

Agents at stage 3 derive utility from the consumption of goods and leisure. The model has two types of goods: a public good and private good. The working-life utility has the following form:

$$u(Q, C, L) = \ln Q + \ln\left(C + \alpha(\ell_{pt} + \ell_{nw}) + \delta \ell_{pt}\right) \tag{3}$$

where $L = (\ell_{ft}, \ell_{pt}, \ell_{nw})$ represents agent's leisure choice, C denotes private consumption and Q corresponds to public consumption of the household. There are three available choices of leisure: full-time employment ℓ_{ft}, part-time employment ℓ_{pt} and not working ℓ_{nw}, such that $\ell_{ft} + \ell_{pt} + \ell_{nw} = 1$. If an agent is a man, then the model limits his choice to two alternatives $\ell_{ft} + \ell_{nw} = 1$, since men outside of training and education rarely actively decide to work part-time (Beham et al., 2019).

Random variable α, which represents a preference for leisure, depends on the agent's marital status $\mathbf{1}\{H^* = \varnothing\}$, gender g and human capital H. Female agents additionally have a preference shifter, denoted by δ, in case they decide to work part-time. Preference shifter δ is a random variable whose values differ for single and married females. Both α and δ are subject to the preferences shocks ξ and υ, respectively. Those shocks are uncorrelated and conditionally on gender follow the normal distribution with zero mean and variance $\sigma^2_{\xi|g}$ and $\sigma^2_{\upsilon|g}$.

Preferences satisfy the *transferable utility* (TU) property if there exists a cardinal representation of utilities, such that for all values of prices and income, the Pareto frontier is a straight line with a slope equal to -1 (Chiappori and Gugl, 2020). One can show that is true for 3 (by taking the $\exp u$ cardinalization). The TU property implies that household aggregate demand does not depend on Pareto weights. It means that at stage 3, a married couple (H, H^*), conditional on labor supply, chooses their optimal consumption of public goods Q and aggregated private consumption $\overline{C}$ $(= C + C^*)$ by solving the following maximization problem:

$$\max_{\overline{C},Q} \exp u(Q, C, L) + \exp u(Q, C^*, L^*) = \max_{\overline{C},Q} Q(\overline{C} + \alpha\ell_{nw} + \alpha^*(\ell^*_{pt} + \ell^*_{nw}) + \delta^*\ell^*_{pt}) \tag{4}$$

with respect to the budget constrain:

$$\overline{Y}^{H,H^*}(L, L^*) \equiv y_{nl}(H, H^*) + \ell_{nw} \cdot b(w) + \ell^*_{nw} \cdot b(w^*) + w_{net}(\ell, \ell^*, w, w^*) = \overline{C} + pQ \tag{5}$$

Household obtains income from work (w, w^*) or unemployment benefits $(b())$. The gross wages are mapped to net income using information about both partners' labor supply and income to mimic a German tax system (details in Appendix E.1). Buettner et al. (2019) provides evidence that households adjust their labor market choices to minimize taxation burden, which makes income mapping an important part of the model. Unemployment benefit $b()$ is defined as a function of wages to mimic the German unemployment benefit system (details in Appendix E.2). Households also obtain a non-labor income (conditional on both partners' human capital), denoted by y_{nl}^{H,H^*}. Household spends the budget on private consumption $\overline{C}$ and public consumption Q. The latter one they buy on the market at a price p.

Conditional on labor supply, the solutions (details in Appendix C.1) for public and private consumption are:

$$pQ(L, L^*) = (\overline{Y}^{H,H^*}(L, L^*) + \alpha\ell_{nw} + \alpha^*\ell^*_{nw} + \delta^*\ell^*_{nw})/2 \tag{6}$$

$$\begin{aligned} \overline{C}(L, L^*) &= \overline{Y}^{H,H^*}(L, L^*) - \alpha\ell_{nw} - \alpha^*\ell^*_{nw} - \delta^*\ell^*_{nw})/2 \\ &= pQ(\ell, \ell^*) - \alpha\ell_{nw} - \alpha^*\ell^*_{nw} - \delta^*\ell^*_{nw}. \end{aligned} \tag{7}$$

Plugging Equations 6 and 7 into the maximization problem given by Equation 4, provides the expression for the optimal choices of labor supply. The final maximization problem is a discrete choice problem. Each couple (H, H^*) has 3×2 choices of labor supply, formally:

$$\max_{L,L^*} pQ^2(L, L^*) \tag{8}$$

The single maximization problem at Stage 3 follows the one presented for couples. Appendix C.2 explains the single maximization problem and describes its solution.

At stage 2, agents do not know the future realization of the productivity and leisure preference shocks. Define $C^* = (\overline{Y}^{H,H^*}(L, L^*) - \alpha\ell_{nw} - \alpha^*(\ell^*_{nw} + \ell^*_{pt}) - \delta^*\ell^*_{pt})/2 - C$, then the ex ante efficient allocation is given by:

$$\max_{C} \mathrm{E}\, u + \mu \,\mathrm{E}\, u^* \tag{9}$$

The solution to this problem is a set of Pareto efficient allocations given by:

$$\begin{aligned} \exp\{\mathrm{E}\, u\}+ \quad & \exp\{\mathrm{E}\, u^*\} = \frac{1}{1+\mu}\exp\{\Psi(H, H^*, r)\} + \frac{\mu}{1+\mu}\exp\{\Psi(H, H^*, r)\} && (10)\\ U_g(H, H^*, r)+ \quad & U_{g^*}(H^*, H, r) = \exp\{\Psi(H, H^*, r)\} \equiv \overline{U}(H, H^*, r) && (11) \end{aligned}$$

where:

$$\Psi(H, H^*, r) \equiv \ln p + \int \ln Q^2(H, H^*, r, \boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}) dF(\boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}) \tag{12}$$

U at stage 3 represents the agent's expected working-life utility from the union (H, H^*) generated at stage 3. Similarly, the function $\overline{U}(H, H^*, r)$ represents the total economic value generated by the couple (H, H^*). It is worth stressing that $\overline{U}(H, H^*, r)$ is the function only of the partners' human capital and region of residence. The detail derivation of both functions are in Appendix C.1).

For single agents, the ex-ante (again, before the realization of the productivity and leisure preference shocks) Pareto efficient set of allocation is defined as:

$$U_g(H, \varnothing, r) = \exp\{\mathrm{E}\, u\} \tag{13}$$

Note that $U_g(H, \varnothing, r)$ refers to the same cardinalisation as in Equation 9.

3.3 Marriage market

At stage 2, agents enter the local marriage markets. They decide whom to marry or to stay single based on preferences and expected utility at stage 3. Let a set of male (female) with human capital H (H^*) living in region r be $N_M^{H,r}$ ($N_F^{H^*,r}$). To identify parameters in the marriage market, I follow the separability assumption in Galichon and Salanié (2021). It states that the total value generated by marriage is a sum of two elements: systematic and idiosyncratic components.

The systematic component consists of an expected economic value obtained by marriage at stage 3 (given by the Equation 9) and taste for similarity (or rather distaste for dissimilarity). The letter one captures the distaste for the divergence in origin (denoted by $\phi_1|o^* - o|$) and the distaste for the difference in education (denoted by $\phi_2|e^* - e|$). Agents (conditional on their human capital and gender) also have a taste for being single.

An idiosyncratic component is the second element of the marriage surplus. Let $\boldsymbol{\omega} = (\omega_{H^*,r} : H^* \in \mathcal{H} \cup \{\varnothing\})$ denote the payoff vector of individual, which represents subjective satisfaction in region r from being married to a person with human capital H^* or staying single. The second part of the separability assumptions stands that an individual draws vector $\boldsymbol{\omega}$ from the probability distribution $\boldsymbol{Q}_g^H$ conditional on gender. It additionally assumes that $\max_{H^* \in \mathcal{H} \cup \{\varnothing\}} |\omega_{H^*,r}|$ have finite expectations under $\boldsymbol{Q}_g^H$.

Formally, the total gain generated by the match between a man with H and a woman H^* living in the region r is:

$$\overline{\Gamma}(H, H^*, r) = \Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) \tag{14}$$

where $\Gamma_M(H, H^*, r)$ and $\Gamma_F(H^*, H, r)$ are partners' individual utilities.

Agents find their preferred partners by maximizing utility. The preferences are characterized by the transferable utility, which means that the surplus given by the Equation 14 is fully divided between spouses. The Pareto weight μ associated with the initial log cardinalization drives the division of the future expected working-life utility. Formally:

$$\Gamma_M(H, H^*, r) = \overline{U}(H, H^*, r) - \tau(H, H^*, r) + \phi_1|o^* - o| + \phi_2|e^* - e| \tag{15}$$
$$\Gamma_F(H^*, H, r) = \tau(H, H^*, r) + \phi_1|o - o^*| + \phi_2|e - e^*| \tag{16}$$

where:

$$\tau(H, H^*, r) = \frac{\mu(H, H^*, r)}{1 + \mu(H, H^*, r)}\overline{U}(H, H^*, r), \quad \mu(H, H^*, r) > 0$$

Pareto weights act as a price that ensures market clearing. This assumption, together with the fact that idiosyncratic shocks are assumed to be independent across two partners, allows me to identify the transfers between agents in the marriage market using marriage outcomes (see Proposition 1 in Galichon and Salanié (2021)).

In the marriage market, some agents match while others do not. A single agent of gender g with human capital H derives utility of the following form:

$$\Gamma_g(\varnothing, r) = U_g(H, \varnothing, r) + \phi_{0H} + \omega_{\varnothing,r} \tag{17}$$

The mapping of who marries whom and who stays single is a match. In the model, I consider the stable match - a match under which no agents have an incentive to deviate from the equilibrium. Formally, the stable match is defined as follows:

Definition 1. A *stable matching* for a marriage market in region r is a triple $(N_M^r, N_F^r, \mathbf{\Gamma}(r))$, where N_M^r (N_F^r) is a set of men (women) living in region r and $\mathbf{\Gamma}(r)$ is a set of payoffs for any men and women, such that for any $H, H^* \in \mathcal{H}$ in r:

1. $\Gamma_M(H, H^*, r) \geq \Gamma_M(H, \varnothing, r)$ for all men
2. $\Gamma_F(H^*, H, r) \geq \Gamma_F(H^*, \varnothing, r)$ for all women
3. $\Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) \geq \overline{\Gamma}(H, H^*, r)$ for all men and women
4. $\Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) = \overline{\Gamma}(H, H^*, r)$ for all matched couples

The first two conditions refer to the *individual rationality* assumption - none of the matched agents can be worse off than while staying single. Condition number 3 refers to the idea of *blocking pairs*. A matching is stable if there are no two agents of the opposite sex such that while matching, they are better off than in their current matching. The last condition states that the sum of individual utilities from marriage equals the total value generated in the match. It is a direct consequence of *transferable utility* assumption.

In theory, Pareto weight μ can be match specific. However, following Chiappori et al. (2018), one can show that μ is specific for a combination partners' human capital (H, H^*). Formally:

Proposition 1. *In a stable match, consider two couples (H, H^*) and (H', H'^*) living in the same region r, such that $H = H'$ and $H^* = H'^*$. Then the Pareto weight is the same for both couples.*
Proof.
From the condition 4 (no blocking pairs) and 5 (transferable utility) of Definition 1 and Equation 14, we have:

$$\begin{aligned}
\Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) &= \overline{U}(H, H^*, r) + 2 \cdot (\phi_1 |e^* - e| + \phi_2 |o^* - o|) + \omega_{H^*,r} + \omega^*_{H,r} \\
\Gamma_M(H, H^*, r) + \Gamma_F(H'^*, H', r) &\geq \overline{U}(H, H'^*, r) + 2 \cdot (\phi_1 |e'^* - e| + \phi_2 |o'^* - o|) + \omega_{H'^*,r} + \omega'^*_{H,r} \\
\Gamma_M(H', H'^*, r) + \Gamma_F(H'^*, H', r) &= \overline{U}(H', H'^*, r) + 2 \cdot (\phi_1 |e'^* - e'| + \phi_2 |o'^* - o'|) + \omega'_{H'^*,r} + \omega'^*_{H',r} \\
\Gamma_M(H', H'^*, r) + \Gamma_F(H^*, H, r) &\geq \overline{U}(H', H^*, r) + 2 \cdot (\phi_1 |e^* - e'| + \phi_2 |o^* - o'|) + \omega'_{H^*,r} + \omega^*_{H',r}
\end{aligned}$$

Then subtracting the first two and the last two equations gives:

$$\omega^*_{H,r} - \omega'^*_{H,r} \geq \Gamma_F(H^*, H, r) - \Gamma_F(H'^*, H', r) \geq \omega^*_{H',r} - \omega'^*_{H',r}$$

which leads to the conclusion:

$$\Gamma_F(H^*, H, r) - \omega^*_{H,r} = \Gamma_F(H'^*, H', r) - \omega'^*_{H,r}$$

It means that the difference between the utility obtained by the wife and her idiosyncratic component is constant across agents with the same human capital. As a result, μ depends only on partners' human capital (H, H^*). $\square$

Proposition 1 shows that, in a stable matching, an individual's utility is simply a sum of three elements: (1) an idiosyncratic shock, (2) noneconomic preferences for partner human capital, (3) endogenously determinate on the labor market share of future economic gain generated by the household. Thanks to that, it is possible to express each individual's problem as a discrete choice problem. It is described in the following proposition:

Proposition 2. *In a stable match, a utility of a man with H satisfies:*

$$\tilde{\Gamma}_M(H, r) = \max_{H^* \in \mathcal{H} \cup \varnothing} \Gamma_M(H, H^*, r) + \omega_{H^*, r} \tag{18}$$

and utility of female j satisfies:

$$\tilde{\Gamma}_F(H^*, r) = \max_{H \in \mathcal{H} \cup \varnothing} \Gamma_F(H^*, H, r) + \omega^*_{H, r} \tag{19}$$

The Proposition 2 states that the discrete choice problem is determined by utility transfers between agents μ, preferences for the difference in partners' education and origin, and individual idiosyncratic shock. μ is exogenous from the agent's perspective. It acts as a price on the marriage market and ensures that no one has an incentive to deviate from stable matching.

Let's assume that agent draws $\boldsymbol{\omega}$ from Extreme Value Type I distribution with variance $\sigma_\omega^{g,H}$. Then, the probability that an agent with H living in the region r marries an agent with human capital H^* is:

$$\mathrm{P}_r\left(H^* | H, r\right) = \frac{\exp\left\{\Gamma_M(H, H^*, r)/\sigma_\omega^{g,H}\right\}}{\sum\limits_{H^* \in \mathcal{H} \cup \varnothing} \exp\left\{\Gamma_M(H, H^*, r)/\sigma_\omega^{g,H}\right\}} \tag{20}$$

At the beginning of Stage 2, agents do not know their idiosyncratic preferences. Using the distribution of $\boldsymbol{\omega}$, the expected utility from stage 2 is given by:

$$\hat{\Gamma}(H, r) = \mathrm{E}\left[\tilde{\Gamma}_M(H, r)\right] = \ln\left(\sum_{H^* \in \mathcal{H} \cup \varnothing} \exp\{\Gamma_M(H, H^*, r)/\sigma_\omega^H\}\right)^{\sigma_\omega^H} + \gamma \tag{21}$$

where γ is an Euler constant.

3.4 Location choice

At Stage 1, agents decide about their future location. There are three possible location choices. Each region is associated with a level of regional amenities, which capture the region's (unrelated to marriage and labor market) attractiveness, e.g., environmental conditions, crime level, transportation system, or general economic situation.

Formally, agents choose their region of residence as follows:

$$r = \arg\max_{r \in \mathcal{R}} \hat{\Gamma}(H, r) + \beta \times Z_r + \eta_r \tag{22}$$

where Z_r represents a vector of regional amenities. $\hat{\Gamma}(H, r)$ is defined like in Equation 21. The choice of the region takes into account both the returns in the labor market and the marriage market structure. Individuals at that point do not know their idiosyncratic components. This assumption corresponds to the situation where agents are unaware of their marital preferences and rather learn about them while meeting new people and dating. η is an idiosyncratic shock, which measures the subjective preferences of agents towards a given region.

If η's are Extreme Type I value distributed with variance σ_η^g, then the probability that agent with H settles down in region r is given by:

$$P(r|H) = \frac{\exp\{(\hat{\Gamma}(H, r) + \beta \times Z_r)/\sigma_\eta^g\}}{\sum_{r' \in \mathcal{R}} \exp\{(\hat{\Gamma}(H, r') + \beta \times Z_{r'})/\sigma_\eta^g\}}$$

Finally, the structure of that stage is a sequential game: agents choose first where they would like to live, but their future utility depends on the distribution of human capital on both sides of the marriage market in the chosen location.

4 Data and Estimation

4.1 Data

This subsection briefly discusses the data and sample used in the estimation. More detailed descriptive statistics of the sample are in Appendix A. There are two primary sources of the data used in this paper. Wages and labor supply choices are estimated using the German Socio-Economic Panel (GSEOP) data for 1984 - 2018. The primary sample includes all males and females aged between 25 and 55[9]. Those enrolled in school or who changed their region of residence are excluded. The final sample contains only singles observed past age 30 to avoid underestimating the marriage rate. For married couples, I only include observations from the first marriage.

Regarding the subsample of immigrants, I exclude those who married before migration since they do not participate in the marriage market in Germany. Additionally, I exclude observations from East Germany due to two reasons. First, the share of immigrants in East Germany is very low, close to zero. Second, economic and law conditions differ in West and East Germany, which can impact household choices. The final dataset contains information on: education (college vs. noncollege), origin (native or immigrant), labor and nonlabor

[9]For couples, the female age and year of birth are reference one.

income, and employment for 94,003 households, of which 78,787 are couples, 6,595 are male singles, and the remaining 8,621 are single women.

The second data source is Microcensus for 2006, 2010, and 2015. The survey collected data with a sampling fraction of 1% of the persons and households in Germany. Due to the size and representativeness, the sample constructed from Microcensus data is used to calculate marriage and region choice probabilities. I divided the sample into 10-year birth cohorts: agents born in the '50s, '60s, and '70s. Then the marriage market and location choice are estimated separately for each cohort.

In the empirical analysis, labor supply decisions are classified into three groups. Full-time workers are those agents who report working at least 35 hours per week. Part-time workers are all female workers who work from 1 to 35 hours per week. All remaining agents are assigned as not-working. The reported wages below the 1st or above the 99th percentile are trimmed to limit the impact of the extreme observation on the estimation result. The model is static, so wages are net of time and age effects. Nonlabor income is replaced, but its estimates net of year and age effects are based on household human capital. Two levels of education correspond to college and noncollege graduates. All agents who were born outside of Germany are qualified as immigrants.

4.2 Outline of the estimation

In this subsection, I discuss the three-step procedure to estimate the model: (1) I estimate outside the model age-year profile and nonlabor income, (2) the estimation proceeds with parameters associated with wages and labor supply choices, (3) the marriage market and location choice parameters are estimated.

4.2.1 Outside the model estimation

I start with an estimation of the nonlabor income of singles and couples on their human capital and region of residence. Predicted nonlabor income is used in the budget constraint at Stage 3. Due to the static form of the model, wages used at Stage 3 are net of age and year effects. It requires estimation of age profile with year fixed effects. I use a control function approach as in Heckman (1979) to allow for endogenous selection to employment. Residuals from the nonlabor income regression and the number of children younger than five are used as exclusion restrictions. To clear wages from age and year effects, I estimate the following equation:

$$\ln \tilde{w} = \ln w + \gamma_0^H age + \gamma_1^H age^2 + \gamma_2^H age^3 + \gamma_4^{year} + \lambda(z^w \beta^w) + \varphi \tag{23}$$

where $\ln w$ is a wage given by Equation 1, γ_4^{year} is a year dummy, and $\lambda(z^w \beta^w)$ is a control function for employment. Vector z^w includes individual residuals from the nonlabor income regression, number of children younger than five, age polynomial, and year dummies. Having estimated the age profile with year dummies, I can replace reported wages with predicted ones. I predict the wages for the age 45 (average age in the sample), keeping their level as in

2005 (sample median). At stage 3, net wages are taken as given. The procedure to estimate wage parameters and leisure preferences is provided below.

4.2.2 Wages and leisure preferences at Stage 3

The utility of a household (H, H^*) at stage 3 is given by the Equation 2. Individuals' preferences for leisure α depend on their human capital and marital status. They are also subject to the leisure preferences shock, formally:

$$\alpha = \alpha_{0g}^{H} + \alpha_{1g}^{H} \cdot \mathbf{1}\{H^* \neq \varnothing\} + \zeta \tag{24}$$

Additionally, females who decide to work part-time have a preference shifter δ. It is a random variable that measures how much female preferences for leisure are different if they decide to work part-time. δ depends on agents' marital status and is subject to the normally distributed shock conditional υ, formally:

$$\delta = \delta_0 + \delta_1 \cdot \mathbf{1}\{H^* = \varnothing\} + \upsilon \tag{25}$$

Wages in the model are as described by Equation 1 and are subject to the productivity shock ε. Productivity and preferences for leisure shocks are drawn from the corresponding distributions at the beginning of Stage 3 (so after the marriage market). It allows me to treat observed marital patterns as given and estimate the postmarital part of the model separately. All wage and leisure preference parameters are estimated using the method of simulated moments. I use 192 moments, which include (1) means, variances, and quantiles of the wage distribution and probability of working and part-time, all by gender, marriage status, and human capital, (2) means and variances of wage distribution by gender, own and partner's human capital to identify the impact of marriages on agents' earnings, (3) means of wage by region, gender and human capital level to identify regional differences in wages. Appendix F contains the complete list of data and simulated moments.

4.2.3 Marriage market and location choices

This subsection briefly describes the estimation procedure for the parameters of the marriage market and location choices. First, I present an estimation strategy for marriage market parameters. Then, I describe the construction of the region amenity index. Finally, I briefly discuss the estimation of location choice parameters.

The marriage market is estimated following the approach by Choo and Siow (2006). In the model, there are two levels of education (college and noncollege) and two origins (native and immigrant). As a result, each agent chooses a partner among four alternatives. In the marriage market, preferences for partners are observable for all participants. Wage shocks and leisure preferences are unknown, so agents match based on expectations. I have estimated in earlier steps the wages and leisure preferences parameters. It allows me to compute the expected economic component for all possible matches defined by the human capital and for singles of all types in each region. I use these estimates to identify the

Pareto weights in this step. The estimation procedure treats Pareto weight and marital preferences parameters as unknown. Marital preferences are identified based on observed choices. Variances of the marital shock are identified based on variations in the expected economic component across regions.

I estimate the model separately for each region and cohort. First, I derive a set of quasi-demand and quasi-supply functions following Choo and Siow (2006) (see Appendix D for details). Since the value of $\overline{U}(H, H^*, r)$ is calculated using estimates from stage 3, the set of Pareto weights is fully identified using the quasi-demand functions (male choices). However, I can also use quasi-supply equations (female choices), which leads to overidentification and allows me to identify the remaining marriage market parameters. I use a minimum distance estimator. The algorithm searches for parameters that minimize the distance between observed quasi-demand and quasi-supply functions and ones implied by the model.

The location choice in the model depends on the region amenities index. The amenity index should ideally capture the whole bundle of amenities, accurately measuring the quality of living in the region. Region amenities index is calculated following the procedure provided by Diamond (2016). First, I collect data on eight different amenities in 16 German regions. The data captures the period from 1971 to 2000, corresponding to the time when cohorts from my sample were aged 21-30. It is a period of life when people are most likely to make their lifetime decision regarding their place of living and marriage. Then, I divide those amenities into four groups: transportation, environmental, crime, and economy. Next, I create amenity subindices using principal component analysis (PCA). Then I use those subindices to calculate an overall amenity index. The final amenity index is aggregated into three regions used in the model using population as a weight.

Table 1 presents the loadings on each amenity subindex and the final overall amenity index. The transportation index negatively weighs the number of passengers but positively length of highways per km^2. It suggests that a single measure of transportation can approximate the development of road networks, which leads to a decrease in the use of public transport. The environmental subindex positively loads the share of forest in the region and the number of national parks. The crime index puts a positive weight on the number of crimes and the number of severe crimes. Finally, the economy index positively weighs GDP per capita and employment, indicating that higher GDP per capita is associated with more jobs.

To create an overall amenity index, I combine all described above subindices. The index accurately places a positive loading on transport, the environment, and the economy. On the other hand, the index weighs negatively on crimes. Intuitively, the amenity index comoves with the safety level. To sum up, a single amenity index constructed based on several subindices represents their common component well.

When choosing the region of residence, agents consider future marriage and labor perspectives. I use estimation parameters from stages 2 and 3 to calculate the continuation value of living in a region for each human capital and cohort. Together with the calculated amenity index, it allows me to estimate parameters associated with location choices. I use the method of moments estimator by minimizing the distance between observed region choice

	Loading	Unexplained variance
Transportation subindex		
Number of passengers in public transport per capita	-0.707	0.220
Length of highways per km^2	0.707	0.220
Environmental subindex		
Forest area in %	0.707	0.162
Number of national parks	0.707	0.162
Crime subindex		
Number of crime cases per capita	0.707	0.282
Number of sever crime cases per capita	0.707	0.282
Economy subindex		
GDP per capita	0.707	0.120
Employment per capita	0.707	0.120
Overall amenity index		
Transport	0.468	0.427
Environmental	0.574	0.138
Crime	-0.550	0.208
Economy	0.386	0.610

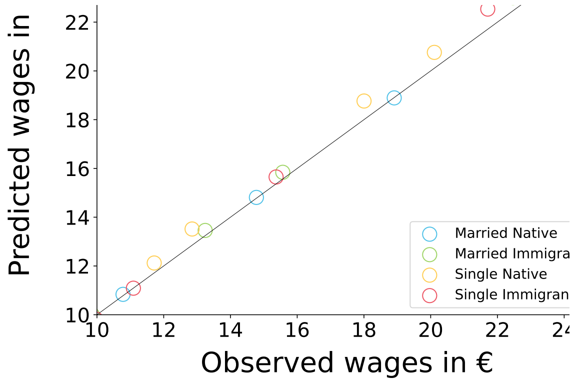
Note: All amenity data measured in standard deviations for the cohort. See Appendix B for detailed description of amenity data.

Table 1: Principle Component Analysis for amenity indices

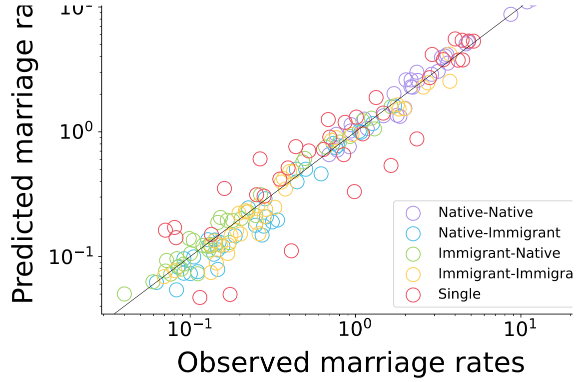
probabilities and the one implied by the model.

4.3 Estimation outcomes

This Section first discuss model's fit. Then, I show estimation of Stage 3, which are those in Equation 1, 24 and 25. Passing then onto a presentation of estimated parameters of Stage 1 and 2 included in 20 and 23.



(a) for wages



(b) for marriage and location

Figure 4: Fit of the model

Notes: Panel A of Figure 4 presents the correlation between observed and predicted wages. Each circle corresponds to a different group defined by marital status, gender, education, and origin. The calculated correlation coefficient is equal to 0.992. Panel B of Figure 4 presents the correlation between observed and predicted marriage rates. The marriage rate is defined as a share of individuals by human capital being married to a type of partner (or staying single) in the total cohort population. The cohorts are agents born in the '50s, '60s, and '70s. The marriage rates are multiplied by 100.

4.3.1 Model fit

The model's fit regarding wages is presented in Figure 4a. The correlation between observed and predicted wages is equal to 0.992. It means that the model quite well explains variation in agents' wages by observable characteristics. The prediction of the model regarding the marriage and location choices are displayed in Figure 4b. The model tends to under- and over-predict the share of single agents, although the general marriage patterns are captured relatively well. The correlation coefficient between observed and predicted marriage rates is equal to 0.989. All targeted moments and their fits for all estimation steps are included in Appendix F.

4.3.2 Wage equations and leisure preferences

Table 2 presents estimated parameters in the wage equation associated with human capital and marriage. In line with the literature findings, there is a positive return to education for all groups of agents. The wages of native agents are higher than immigrants with the same education level. It can suggest that human capital is not fully transferable between countries. When it comes to parameters associated with marriage, they are positive for married men (except for college-educated immigrants) and negative for women (except for noncollege immigrants). It would suggest that women are penalized in the labor market when married. In some cases, the value of the marriage premium depends on the partner's origin. However, the effect is not strong (coefficients are insignificant at a confidence level of 0.05), which could suggest that the partner's origin impacts total household labor income but does not influence the agent's wage.

	Native noncollege	Native college	Immigrant noncollege	Immigrant college
Male				
Constant	2.378	2.817	2.140	2.653
	(0.017)	(0.026)	(0.054)	(0.089)
Married	0.178	0.210	0.214	-0.000
	(0.012)	(0.022)	(0.043)	(0.094)
Married to immigrant	0.006	0.050	0.006	-0.012
	(0.019)	(0.037)	(0.028)	(0.080)
Female				
Constant	2.319	2.813	1.987	2.436
	(0.026)	(0.021)	(0.060)	(0.089)
Married	-0.146	-0.057	0.054	-0.071
	(0.013)	(0.019)	(0.043)	(0.076)
Married to immigrant	0.017	0.029	-0.034	0.058
	(0.027)	(0.056)	(0.033)	(0.085)

Notes: Standard errors in parentheses. Wages are expressed in log values, deflated using 2005 prices. Base category: single living in North.

Table 2: Wage parameters - human capital and marriage premium

These estimates allow the model to predict some regularities in wages, migration, and marriages established in the literature. Returns to education are quantitatively similar to those presented in Card (1999). Psacharopoulos and Patrinos (2004) finds that, on average, females have higher education returns than males. The estimated parameters of the model are in line with this finding when it comes to the married population. I find lower returns to education among immigrants confirm the less-than-perfect international transferability of human capital found in, i.e., Chiswick and Miller (2009). The effect of intermarriage on wages is positive but insignificant, which supports the hypothesis by, i.e., Kantarevic (2004) and is contrary to the finding of Meng and Gregory (2005).

Estimation outcomes of wage equations have significant implications for the marriage market. The estimated parameters suggest that when immigrants decide to marry a native compared to immigrants, they can count on higher household income in the future. On the other hand, marriage with an immigrant does not pay off for natives - mixed household labor income could be lower than in the case of all-native households. The final size of the economic gain of immigrants from a marriage with a native depends on their bargaining power within the household. A higher Pareto weight could compensate for the lower household income. Shifting part of the income from immigrants to natives increases the relative attractiveness of immigrants from a native perspective.

Table 3 contains outcomes of the region fixed effect estimation in wage equation. First, on average, agents earn the highest wage in the South region and the lowest in the North. Second, returns to human capital are not homogeneous across regions, and immigrants experience stronger variations. Noncollege immigrants earn more in the South than in the North on average by 13.4%. The effect for college-educated immigrants is even stronger and equals 28.3%. It could suggest that the South is an attractive migration destination for foreign-born individuals. Agents who live in region West also earn more than those living in region North. However, the positive effect is weaker (3.2% vs. 13.4% for noncollege immigrants) and more homogeneous than in the region South (the difference between estimated coefficients is smaller).

	Human capital			
Region	Native noncollege	Native college	Immigrant noncollege	Immigrant college
South	0.075	0.090	0.134	0.283
	(0.010)	(0.017)	(0.032)	(0.084)
West	0.046	0.075	0.032	0.138
	(0.011)	(0.017)	(0.033)	(0.083)

Notes: Standard errors in parentheses. Wages are expressed in log values, deflated using 2005 prices. Base category: North.

Table 3: Wage parameters - differences across regions

I present estimated parameters associated with leisure preferences in Table 4. Single males have, on average lower leisure preferences than married ones for all human capital types, while the opposite is true for females. This difference can be partially explained by a higher number of children among married couples in comparison to singles. Higher fertility may cause married men to take a job more often, while women drop from the labor market to take care of children.

Interestingly, on average, immigrants have higher leisure preferences than natives. It could be related to several things. First, they may produce more at home. Home production is not included in the model so it may be partially captured by the parameter α. Home production is also more important for immigrants, since they may face problems to buy ethnic products on the market. Second, immigrants in the data have, on average, more children than natives, which can also contribute to higher leisure preferences. Part-time work shifter δ is positive for females and higher for single ones. It could suggest that in this group, mixing work with leisure (potentially taking care of children) is an additional source of utility.

Results presented in Table 2 and 4 suggest significant differences between natives and immigrants regarding wages and leisure preferences exist. Those differences are demonstrated in the economic value generated by single households at Stage 3. Table 5 presents it for

	Male		Female	
	constant	married	constant	married
Pref. for leisure, α				
Native noncollege	1.591	-1.859	0.097	1.690
	(0.117)	(0.167)	(0.394)	(0.626)
Native college	0.150	-1.759	-1.451	1.818
	(0.347)	(0.301)	(0.660)	(0.691)
Immigrant noncollege	2.974	-1.574	1.706	1.877
	(1.410)	(1.404)	(0.329)	(0.489)
Immigrant college	2.776	-0.589	0.801	2.893
	(0.867)	(0.936)	(0.633)	(0.850)
Part-time work shifter, δ			6.217	-5.335
			(1.799)	(1.809)

Notes: Standard errors in parentheses. Column 'married' presentes change in parameters value for married.

Table 4: Leisure parameters - preferences for not-woring and part-time working

singles living in region West. Economic value is increasing in education. It is also higher for natives than for immigrants. Similar patterns are found in Table 6, which presents the expected marriage economic value for couples living in the region West for all 16 possible combinations of human capital. There are two important conclusions from this. First, from a female point of view, marriage with a native generates higher economic gain than marriage with an immigrant with the same education level. The reason is that most men work (so their leisure preferences matters less), and native men have higher wages than an immigrant. Second, from an immigrant male point of view, marriage with a native also generates higher economic gain than marriage with an immigrant with the same education level. Interestingly, it is not true for native men. The reason could be that immigrant women have lower wages and higher preferences for leisure, so when they marry a native, they more often do not work and generate more utility from this choice than native women.

	Human capital			
	Native noncollege	Native college	Immigrant noncollege	Immigrant college
Male	24.81	51.79	16.38	49.06
Female	31.41	49.45	29.98	38.78

Table 5: Economic value of staying single in region West

Male human capital	Female human capital			
	Native noncollege	Native college	Immigrant noncollege	Immigrant college
Native noncollege	76.65	96.81	79.21	98.42
Native college	134.31	158.57	151.37	176.05
Immigrant noncollege	69.36	92.83	67.47	92.72
Immigrant college	101.11	129.73	97.35	128.09

Table 6: Economic value of marriage - West

4.4 Marriage market and location choice

This subsection briefly describes the estimation outcomes of parameters associated with the marriage market and location choice. Table 7 contains estimated parameters for marital preferences. The parameters are estimated separately for each birth cohort - agents born in the '50s, '60s, and '70s of the XX century. The left-hand side of the table presents estimates of preferences for being single. They are higher for college-educated agents and natives. Also, preferences for being a single increase over time, which is in line with the data that suggest that share of people who decide to stay single is increasing. The right-hand side of Table 7 contains estimated parameters associated with a taste for dissimilarity in education and origin. Estimates are negative for differences in education level and origin, but the distaste for dissimilarity in origin is higher. It could be correlated with higher social norms, which need to be broken when one marries a person from a different origin group. Interestingly, the interaction term is positive and offsets the negative effect of a difference in origin and education in some parts. It could suggest that breaking both norms is associated with smaller negative tastes. All these estimates suggest that the agent prefers to marry people of the same origin and education. As a result, they have a strong tendency to trade better economic outcomes for similarities. It leads to highly positive assortative mating in the marriage market.

Agents choose a partner based on individual marital preferences and the share of the economic gain generated by the couple, which corresponds to the Pareto weight of the collective household model. The share is unique for each type of couple living in a region. It depends not only on the agent's human capital but also reflects the relative scarcity of spouses. Therefore, it depends on the entire human population distribution in the given region. The share acts as a price that clears the marriage market. Table 8 presents a share of the gains from a marriage that belongs to women for couples born in the '60s and living in the West region. A higher education level generally correlates with a higher share of future utility. If a college-educated immigrant woman marries a noncollege immigrant, she gets 81% of welfare. Similarly, if a college-educated native man wants to marry a college-educated native woman, her share will be higher than that of noncollege-educated native women by around 20 pp. When the patterns for education are clear, it is not valid for origin. On average,

ϕ_{0H}	Native noncollege	Native college	Immigrant noncollege	Immigrant college	$\phi_1\lvert e-e^*\rvert$	$\phi_2\lvert o-o^*\rvert$	$\phi_3\lvert e-e^*\rvert\cdot\lvert o-o^*\rvert$
'50	10.220	65.184	10.622	49.522	-20.042	-28.990	3.466
	(2.313)	(4.565)	(6.629)	(4.688)	(2.224)	(2.328)	(0.793)
'60	17.575	69.070	14.387	52.878	-18.381	-28.780	4.449
	(1.941)	(4.635)	(6.438)	(4.636)	(2.127)	(2.390)	(1.041)
'70	20.851	73.082	16.356	58.155	-18.260	-25.045	2.888
	(1.795)	(4.613)	(6.349)	(4.355)	(1.996)	(2.116)	(0.671)

Notes: Asymptotic bootstrapped standard errors in parentheses. The left-hand side of the table presents estimates of the taste for staying single by cohort and human capital. The right-hand side of the table presents estimates of taste for similarity by cohort.

Table 7: Marriage market parameters - taste for staying single and similarity

women can extract a higher share from native men than immigrants of the same education level. It is also true for men regarding native and immigrant women. It reflects the scarcity of immigrants in comparison to natives. The higher share attributed to immigrant women compared to immigrant men suggests that immigrant women have better opportunities and stronger bargaining power while intermarrying than immigrant men.

	Female human capital			
Male human capital	Native noncollege	Native college	Immigrant noncollege	Immigrant college
Native noncollege	0.567	0.797	0.698	0.859
	(0.015)	(0.009)	(0.012)	(0.021)
Native college	0.330	0.522	0.355	0.489
	(0.009)	(0.009)	(0.023)	(0.017)
Immigrant noncollege	0.484	0.685	0.661	0.805
	(0.013)	(0.032)	(0.024)	(0.035)
Immigrant college	0.218	0.440	0.330	0.516
	(0.028)	(0.021)	(0.032)	(0.030)

Notes: Asymptotic standard errors in parentheses are computed using the bootstrap method.

Table 8: Sharing rule among couples born in '60s, living in region West

Table 9 presents estimation of parameters associated with location choice. The parameter associated with region amenities is positive for all human capital levels. It is the strongest for noncollege immigrants. It suggests that regional amenities play the highest role for this group when they choose their future place of living. It causes agents with a higher taste for amenities to be less sensitive to changes in wages and marriage market conditions.

	Native noncollege	Native college	Immigrant noncollege	Immigrant college
β	0.106	0.089	0.174	0.049
	(0.004)	(0.008)	(0.006)	(0.020)

Notes: Standard errors in parentheses.

Table 9: Taste for amenities

5 Policy scenarios

This section presents the analysis of two policy scenarios. First, I briefly describe the scenarios. Next, I present changes to single and intermarriage rates under each scenario and discuss the consequences of those changes for the integration of immigrants. Further, I show how immigrants' concentration transforms and how it impacts integration. Finally, I present income inequality and welfare changes in each case with and without adjustment in marriage and location choices.

Policymakers might target immigrants' integration through various labor market policies. Reducing the immigrant-native wage gap is indispensable to holistic integration (Lehmer and Ludsteck, 2015). In scenario I, the government introduces a policy that directly impacts immigrants' wages. As a result, the immigrant-native wage gaps across genders and all education levels decrease by 25%. It is an equivalent to an average increase in wages of immigrant men by 4.6% and of immigrant women by 7.5%. Instead of a universal closer to the wage gap, a government might prefer to focus on regional wage variation. In scenario II, the introduced policy leads to the closing of the difference between the best region (South) and the remaining regions (North and West) by increasing the average wage in the remaining regions by 50% of the initial size of the gap. This change is equivalent to an average increase in immigrants' wages by 3.5%. Both scenarios improve immigrants' labor market outcomes, which might change their marital and spatial distribution. Immigrants, who get richer, become more attractive to natives, and the intermarriage rate might increase. On the other hand, the relative gain from marrying a native decreases from the immigrant perspective, which can lead to the opposite outcome. Similarly, adjustments in labor and marriage market conditions could influence overall region attractiveness and lead to changes in the spatial segregation of immigrants. The direction and magnitude of changes are evaluated using the model and estimated parameters presented in the previous chapter of the paper.

Table 10 presents changes in the spatial concentration of immigrants for both scenarios by gender and education level. The spatial concentration is measured as a total variation distance between uniform and observed distribution. Distribution across regions of women and college-educated agents seems to be more concentrated, but the differences are relatively minor. Increasing the wages of immigrants such that the gap between them and natives drop by 25% lead to a decrease in the spatial distribution of all groups. The size of the effect is limited, however, it might improve immigrants' integration. The decline is more substantial among women, especially the college-educated ones. It comes from the fact that the initial wage gap among immigrant women is the biggest, so they respond stronger to changes in its relative value. The decrease in the spatial concentration of immigrants is much stronger in the case of scenario II. The effect is massive among college-educated immigrants and equals around 32% for both genders. It suggests that decreasing the regional wage gap among immigrants has a strong and positive effect on their immigration, measured by spatial concentration.

	Baseline	Scenario	
		I	II
		Δ%	Δ%
A. Men			
Noncollege	0.132	-1.7%	-9.3%
College	0.143	-1.7%	-31.7%
Total	0.134	-1.7%	-15.1%
B. Women			
Noncollege	0.146	-2.9%	-11.8%
College	0.147	-8.5%	-31.9%
Total	0.146	-4.0%	-15.9%

Notes: Cells in the table present the spatial concentration of immigrants and its changes under counterfactual scenarios. The spatial concentration of immigrants is measured using the total variation distance between uniform and observed distributions. The sample consists of a cohort of agents born in the '60s.

Table 10: Changes in immigrants' spatial concentration

Analyzed scenarios improve the economic situation of immigrants. Further, it induces changes in marriage market outcomes since matching depends on, among others, economic value. Table 11 presents changes to the single rate among immigrant men and women by education level. The first column contains a prediction of the model based on estimated consistent parameters. The remaining columns present the percentage point differences between the baseline for each scenario. Under scenario I (reduction in immigrant-native wage gap), the share of single immigrant decrease in all groups (from -0.2% to -8.2%). The decrease is more substantial among men than among women. It occurs because men work, on average, more, so wage changes impact them to a greater extent. Also, college-educated immigrants react stronger than noncollege ones (-8.2% vs -2% and -1.9% vs -0.2%). The discrepancy, again, is driven by differences in labor supply. Similar patterns can be observed under the scenario I, but the decrease in the single rate is smaller, except for noncollege educate women. A decrease in the single rate suggests that considered labor market changes increase immigrant attractiveness as partners and induce more marriages. Understanding the effect of these changes on social integration requires an analysis of changes in marriage patterns.

	Baseline	Scenario	
		I	II
	%	Δ	Δ
A. Men			
Noncollege	18.5%	-2.0%	-1.6%
College	14.8%	-8.2%	-5.3%
Total	17.8%	-3.2%	-2.3%
B. Women			
Noncollege	5.3%	-0.2%	-0.5%
College	9.0%	-1.9%	-1.8%
Total	6.1%	-0.5%	-0.8%

Notes: Cells in the table present the shares of single immigrantsand their percentage point changes under counterfactual scenarios. The share of single immigrants is calculated as the number of singles divided by the total population by gender, origin, and education group. The sample consists of a cohort of agents born in the '60s.

Table 11: Single rate among immigrants

Table 12 presents the share of immigrants married to natives by gender and education level. The first column contains model prediction using consistently estimated parameters. The remaining columns present the percentage point differences between the baseline and two counterfactual scenarios. Panel A of Table 12 contains those values for immigrant men. Under both scenarios, the share of intermarried male immigrant change in a very similar way. Introduced policies increase the probability of being married to a native. However, the increase in the intermarriage rate is much lower than the increase in the marriage rate showed in Table 11, indicating that single immigrants in the baseline scenario still more often marry other immigrants than natives in the alternative scenarios. However, the overall effect of both policies on the social integration of immigrant men is positive.

Panel B of Table 12 presents the same values as panel A but for immigrant women instead. In scenario I, the share of intermarried women decreases by 1.9 p.p. for noncollege- and 2.4 pp for college-educated. It occurs because the utility of marriage with male immigrants increases (due to wage growth caused by closing the immigrant-native gap), and they become more attractive partners. The difference between the two education groups comes from the fact that college-educated immigrant women are more likely to be single than noncollege ones in the baseline scenario. It means that the decrease in intermarriage comes not only from women who change their partners' origin. It also comes from women who are single under the baseline scenario but, due to policy changes, not anymore, and they marry immigrant men more often. In the case of scenario II, the magnitude of changes is smaller (-0.5 pp and -1 pp, respectively). However, the relative value and sources of the changes remain the same. It means that both policies negatively impact the social integration of immigrant women.

	Baseline	Scenario	
		I	II
	%	Δ	Δ
A. Men			
Noncollege	28.8%	0.9%	0.9%
College	32.3%	2.2%	2.1%
Total	29.5%	1.1%	1.1%
B. Women			
Noncollege	31.3%	-1.9%	-0.5%
College	32.4%	-2.4%	-1.0%
Total	31.5%	-2.0%	-0.6%

Notes: Cells in the table present the shares of intermarried immigrants and their percentage point changes under counterfactual scenarios. The share of intermarried immigrants is calculated as the number of immigrants married to natives divided by the total population of agents by gender, origin, and education group. The sample consists of a cohort of agents born in the '60s.

Table 12: Intermarriage rate among immigrants

Despite the integration aspects, such as the intermarriage rate or spatial segregation of immigrants, policymakers also care about the consequences of their policies on income inequality or welfare. Table 13 presents the baseline level of income inequality and its percentage point changes under two scenarios with and without adjustments for location and marriage choices. Income inequality is calculated as a percentage difference between the average per capita income of immigrants and natives by gender and education level. The baseline values are slightly more unfavorable for women than for men. Reducing the immigrant-native wage gap decreases income inequality (from 3.2% to 1.9%) without controlling for adjustments. Even though the average wage increase is higher for women than men due to labor supply choices, the positive effect on income inequality is weaker for women than for men. Further, allowing for adjustments in marriage and region choices leads to a decrease in the effect size. In the case of college-educated men, the primary source of the decrease is that under the counterfactual scenario, they are more often married, so they share the income with partners who are less likely to work. In the case of women, the decrease comes from the fact that under the baseline scenario, they are more often married to natives. The better labor situation of immigrant men makes immigrant women willing to trade higher income for taste in similarity, which mitigates the positive effect on income inequality. The exception to this pattern are noncollege-educated men. In their case, the income inequality gets even smaller while accounting for the marriage market adjustments. It comes from the fact that they are the only group that marries more often with natives, who have higher wages and work more often. Regarding scenario II, the changes in income inequality are smaller but stay positive. They also show similar patterns when controlling or not for adjustments in the marriage market and region choices. Analyses of counterfactual scenarios suggest that both outcomes of labor market integration politics decrease income inequality. However, changes in region and partner choices partially mitigate the positive effect for all groups except noncollege-educated men.

	Baseline	Scenario I			Scenario II		
		(1)	(2)	(3)	(1)	(2)	(3)
		Δ	Δ	Δ	Δ	Δ	Δ
A. Men							
Noncollege	-12.6%	2.4%	2.7%	2.7%	1.7%	1.9%	1.9%
College	-11.3%	3.2%	2.5%	2.5%	3.0%	2.4%	2.1%
B. Women							
Noncollege	-13.2%	1.9%	1.7%	1.6%	1.4%	1.4%	1.3%
College	-14.5%	3.0%	3.0%	2.9%	2.3%	2.5%	2.0%
Marriage market adj.		×	✓	✓	×	✓	✓
Region choice adj.		×	×	✓	×	×	✓

Notes: Cells in table present income inequality and its percentage point changes under counterfactual scenarios. Income inequality measured as parecentage difference between average per capita income of immigrant and native of the same gender and education level.The sample consists of a cohort of agents born in the '60s.

Table 13: Changes in income inequality

Table 14 presents welfare changes under two policy scenarios. The changes are disaggregated in the same way as in Table 13. In scenario I, there is an increase in the welfare of all immigrants (from 1.2% to 3.2%), keeping the baseline distribution of marriages and location choices. Since immigrants' wages rise, households with immigrants have higher disposable income, which leads to utility gains. The increase is more significant for college-educated immigrants than for noncollege-educated ones. The difference is driven by the initial higher wages of better-educated agents. Allowing for adjustment in marital choices increases the welfare of all immigrants except noncollege-educated women. The less positive or negative impact of marriage market adjustment for female welfare emerges from the bargaining between partners. Men work, on average, more than women, so increasing their wages improves their negotiation situation. As a result, they can negotiate more favorable Pareto weights. Endogenizing regional choices do not significantly change welfare outcomes because closing the immigrant-native gap is parallel in all regions. The welfare consequences of reducing regional variation in immigrants' wages are positive for all types of immigrants. The patterns of changes are similar to the one under scenario I, except for positive change induced by adjustment in region choice. The increasing welfare gains associated with adjustment in marriage and location choices contradict rising income inequality shown in Table 13. It suggests that even though the income per capita after adjustment decreases, it is compensated by the higher utility of marriage. In particular, agents compensate for the decreasing income by a taste for similarity, marrying more often with other immigrants.

	Scenario I			Scenario II		
	(1)	(2)	(3)	(1)	(2)	(3)
A. Men						
Noncollege	1.2%	1.8%	1.8%	0.8%	1.1%	1.1%
College	3.2%	3.7%	3.7%	2.4%	2.7%	2.9%
Total	1.7%	2.3%	2.3%	1.3%	1.5%	1.6%
B. Women						
Noncollege	1.5%	1.4%	1.4%	1.1%	1.1%	1.1%
College	2.8%	3.0%	3.0%	2.0%	2.4%	2.5%
Total	1.8%	1.8%	1.8%	1.3%	1.4%	1.5%
Marriage market adj.	✗	✓	✓	✗	✓	✓
Region choice adj.	✗	✗	✓	✗	✗	✓

Notes: Cells in table present percentage changes in welfare of immigrant men and women by human capital.The sample consists of a cohort of agents born in the '60s.

Table 14: Welfare changes for immigrants

An analysis of the two policy scenarios reveals that changes in labor market conditions lead to an adjustment in marital and spatial distribution. It underlines the importance of including those two aspects in the analysis of outcomes of labor market integration policies. First, analyzed policies decrease the spatial concentration of immigrants, which could strengthen the positive effect of labor market changes for integration. Second, both policies increase the number of marriages and intermarriages among immigrant men. They also increase the number of marriages among immigrant women but decrease intermarriages among them. It suggests that improving immigrants' labor market conditions differently impacts the social integration of men and women via marriages with natives. Third, income inequality increases while I control for adjustment in marriage and location choices. It suggests that immigrants trade economic gains for better region or marriage perspectives. Finally, counterfactual scenarios showed that improving the labor market situation of immigrants allows for obtaining additional welfare gains through marital and regional sorting changes.

6 Conclusion

In this paper, I estimate a three-stage structural model using German micro-data to quantify the effect of labor market integration policies on intermarriage and spatial concentration of immigrants. Further, I evaluate how those changes impact immigrant-native income inequality and immigrants' welfare. Estimated parameters suggest that immigrants, on average, earn less than natives of the same gender and education. It makes them less attractive partners and, together with estimated strong preferences for similarity, leads to a relatively low number of intermarriages. I use the model to simulate the effects of reducing the immigrant-native wage gap by 25% and decreasing regional variation in immigrants' wages by 50%. These exercises lead to three main conclusions. First, closing the immigrant-native wage gap positively affects the spatial concentration of immigrants because regions with a more significant gap (and, at the same time, smaller migration populations) become more attractive. Further, the policy's impact on the frequency of marriages with natives varies by immigrant gender. The increase in wages stronger affects the economic attractiveness of immigrant men since they have a higher labor supply. As a result, the policy leads to more marriages between them and native women. On the other hand, due to the gender difference in the size of the policy's effect, immigrant women marry immigrants more often. As a result, the changes in homogamy differently affect the social integration of immigrant men and women. Second, reducing regional variation in immigrants' wages significantly changes their regional distribution. Immigrants move out from the South (the region offering the highest wages in the baseline scenario) to settle in the North and West. It significantly flattens the distribution and leads to a decrease in spatial concentration across genders and education levels up to 32%. This finding suggests that reducing regional wage variation by increasing immigrants' earnings in regions initially characterized by lower wages might be a powerful policy tool. Not only it decreases income inequalities, but also positively impacts integration via a decline in spatial concentration. However, these positive boosts are partially mitigated by the decrease in the intermarriage rate among immigrant women. Finally, both policy scenarios improve welfare levels and lower the income inequality between immigrants and natives. Adjustments in marriage and location choices lead to, on average, a simultaneous decrease in the drop of income distribution and an increase in welfare gains. It suggests that after improvement in labor market conditions, immigrants trade part of the gains for noneconomic gains associated with marriage and location choices.

References

Adda, J., Pinotti, P., and Tura, G. (2020). There's More to Marriage than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations. *CEPR Discussion Papers*, (November).

Basu, S. (2015). Intermarriage and the labor market outcomes of Asian women. *Economic Inquiry*, 53(4):1718–1734.

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

Becker, G. S. (1974). A Theory of Marriage: Part II. *Journal of Political Economy*, 82(2, Part 2):S11–S26.

Beham, B., Drobnič, S., Präg, P., Baierl, A., and Eckner, J. (2019). Part-time work and gender inequality in Europe: a comparative analysis of satisfaction with work–life balance. *European Societies*, 21(3):378–402.

Benson, S. (1981). *Ambiguous Ethnicity: Inter-racial Families in London.* Cambridge University Press, Cambridge.

Boeri, T., De Philippis, M., Patacchini, E., and Pellizzari, M. (2015). Immigration, Housing Discrimination and Employment. *Economic Journal*, 125(586):F82–F114.

Borjas, G. J. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics*, 3(4):463–489.

Buettner, T., Erbe, K., and Grimm, V. (2019). Tax planning of married couples and intra-household income inequality. *Journal of Public Economics*, 179(3):104048.

Calvo, P. A., Lindenlaub, I., and Reynoso, A. (2021). Marriage Market and Labor Market Sorting. *Ssrn*, (July).

Card, D. (1999). Chapter 30 The causal effect of education on earnings. *Handbook of Labor Economics*, 3 PART(1):1801–1863.

Chiappori, P. A., Dias, M. C., and Meghir, C. (2018). The marriage market, labor supply, and education choice. *Journal of Political Economy*, 126:S26–S72.

Chiappori, P.-A. and Gugl, E. (2020). Transferable utility and demand functions. *Theoretical Economics*, 15(4):1307–1333.

Chiappori, P. A., Salanié, B., and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8):2109–2167.

Chiswick, B. R. and Houseworth, C. (2011). Ethnic intermarriage among immigrants: Human capital and assortative mating. *Review of Economics of the Household*, 9(2):149–180.

Chiswick, B. R. and Miller, P. W. (2009). The international transferability of immigrants' human capital. *Economics of Education Review*, 28(2):162–169.

Choi, K. H. and Tienda, M. (2017). Marriage-Market Constraints and Mate-Selection Behavior: Racial, Ethnic, and Gender Differences in Intermarriage. *Journal of Marriage and Family*, 79(2):301–317.

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1):175–202.

Cohen-Goldner, S. and Eckstein, Z. (2008). Labor mobility of immigrants: Training, experience, language, and opportunities. *International Economic Review*, 49(3):837–872.

Cohen-Goldner, S. and Eckstein, Z. (2010). Estimating the return to training and occupational experience: The case of female immigrants. *Journal of Econometrics*, 156(1):86–105.

Compton, J. and Pollak, R. A. (2007). Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas? *Journal of Labor Economics*, 25(3):475–512.

Costa, D. L. and Kahn, M. E. (2000). Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990. *The Quarterly Journal of Economics*, 115(4):1287–1315.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (2008). Is the melting pot still hot? Explaining the resurgence of immigrant segregation. *Review of Economics and Statistics*, 90(3):478–497.

Danzer, A. M. and Yaman, F. (2013). Do ethnic enclaves impede immigrants' integration? Evidence from a quasi-experimental social-interaction approach. *Review of International Economics*, 21(2):311–325.

Diamond, R. (2016). The determinants and welfare implications of US Workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.

Dribe, M. and Lundh, C. (2011). Cultural dissimilarity and intermarriage. A longitudinal study of immigrants in Sweden 1990-2005. *International Migration Review*, 45(2):297–324.

Elwert, A. (2018). *Will You Intermarry Me? : Determinants and Consequences of Immigrant-Native Intermarriage in Contemporary Nordic Settings.*

Elwert, A. and Tegunimataka, A. (2016). Cohabitation Premiums in Denmark: Income Effects in Immigrant–Native Partnerships. *European Sociological Review*, 32(3):383–402.

Fan, J. and Zou, B. (2021). The Dual Local Markets: Family, Jobs, and the Spatial Distribution of Skills. *SSRN Electronic Journal*, (November).

Galichon, A. and Salanié, B. (2021). Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *The Review of Economic Studies*, (December):1–30.

Grossbard, S. A. and Vernon, V. (2020). Do immigrants pay a price when marrying natives? Lessons from the US time use survey. *IZA Journal of Development and Migration*, 11(1).

Harris, D. R. and Ono, H. (2005). How many interracial marriages would there be if all groups were of equal size in all places? A new look at national estimates of interracial marriage. *Social Science Research*, 34(1):236–251.

Hayfron, J. E. (2001). Language training, language proficiency and earnings of immigrants in Norway. *Applied Economics*, 33(15):1971–1979.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Joona, P. A. and Nekby, L. (2012). Intensive Coaching of New Immigrants: An Evaluation Based on Random Program Assignment. *Scandinavian Journal of Economics*, 114(2):575–600.

Kalmijn, M. and van Tubergen, F. (2006). Ethnic intermarriage in the Netherlands: confirmations and refutations of accepted insights. *European Journal of Population / Revue européenne de Démographie*, 22(4):371–397.

Kantarevic, J. (2004). Interethnic Marriages and Economic Assimilation of Immigrants. *IZA Discussion Paper*, (1142).

Lazear, E. P. (1999). Culture and language. In *Journal of Political Economy*, volume 107, pages S95–S126.

Lehmer, F. and Ludsteck, J. (2015). Wage Assimilation of Foreigners: Which Factors Close the Gap? Evidence From Germany. *Review of Income and Wealth*, 61(4):677–701.

Llull, J. (2018). The Effect of Immigration on Wages. *Journal of Human Resources*, 53(3):608–662.

Lochmann, A., Rapoport, H., and Speciale, B. (2019). The effect of language training on immigrants' economic integration: Empirical evidence from France. *European Economic Review*, 113:265–296.

Meng, X. and Gregory, R. G. (2005). Intermarriage and the economic assimilation of immigrants. *Journal of Labor Economics*, 23(1):135–175.

Meng, X. and Meurs, D. (2009). Intermarriage, language, and economic assimilation process: A case study of France. *International Journal of Manpower*, 30(1-2):127–144.

Min Zhou and Logan, J. R. (1989). Returns on human capital in ethnic enclaves: New York City's Chinatown. *American Sociological Review*, 54(5):809–820.

OECD (2020). *International Migration Outlook 2020*. OECD.

Psacharopoulos, G. and Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, 12(2):111–134.

Roback, J. (1982). Wages , Rents , and the Quality of Life Jennifer Roback. *Journal of Political Economy*, 90(6):1257–1278.

Rodríguez-García, D. (2015). Intermarriage and Integration Revisited: International Experiences and Cross-Disciplinary Approaches. *Annals of the American Academy of Political and Social Science*, 662(1):8–36.

Rosen, K. T. (1979). A Regional Model of Multifamily Housing Starts.

van Tubergen, F. and Maas, I. (2007). Ethnic intermarriage among immigrants in the Netherlands: An analysis of population data. *Social Science Research*, 36(3):1065–1086.

Xie, Y. and Gough, M. (2011). Ethnic Enclaves and the Earnings of Immigrants. *Demography*, 48(4):1293–1315.

A Descriptive statistics

Table 15 reports the descriptive statistics for a subsample of immigrants by gender and three marriage statuses: single, inter- and noninter- married. The mean age and standard deviation for intermarried and nonintermarried immigrants are similar for males. The sample of intermarried females seems to be slightly older than nonintermarried ones. Intermarried male and female immigrants, on average, spend more time in Germany by around half a year compared to male immigrants who are not married to German women. Married male immigrants, on average, migrate at the age of 19. In the case of women, intermarried immigrants arrive in the host country at 20, two years older than nonintermarried ones. The difference in years of education between intermarried and nonintermarried immigrants is the same for males and women and equals one year.

The second part of Table 15 presents statistics associated with immigrants' assimilation. On average, female immigrants married to Germans declare that they feel more German and less often that they do not belong to German society than nonintermarried immigrants. For male immigrants, the relationship is the same regarding feeling that they do not belong to German society. Intermarried immigrants are also characterised by better, on average, knowledge of oral German, and they more often use German media. Immigrants married to other immigrants report being visited by German less frequently than intermarried ones. This data suggest that intermarried immigrants are, on average, better assimilated than nonintermarried ones.

B Data sources

Table 16: Amenity Indices - data sources

Variable	Source	Sample	Notes
Population	Polizeiliche Kriminalstatistik 1966, 1976, 1986, 1996, 206, 2016	All Bundelands	
Number of passengers in public transport	Statistisches Jahrbuch für die Bundesrepublik Deutschland 1965, 1976, 1986, 1995; Destatis, Tabelle 46181-0011	All public companies	Total number of passengers transported by public companies within calendar year
Length of highways	Statistisches Jahrbuch für die Bundesrepublik Deutschland 1966, 1976, 1986, 1997, 2005; Statistisches Jahrbuch 2015	All public roads	Total length of highways available at the end of the calendar year

Number of national parks	Bundesamt für Naturschutz, Naturparke in Deutschland (01/01/2020)	All national and natural parks	If national parks is a part of more than one Bundesland, then it was assigned to all of them.
Forest area in %	Fachserie B. Land- und Forstwirtschaft, Fischerei. Statistisches Bundesamt Wiesbaden. 1964, 1974, 1985, 1993; Tabelle 33111-0004: Bodenfläche (tatsächliche Nutzung): Bundesländer, Stichtag (bis 31.12.2015), Nutzungsarten	Area classified as a forest	The raw area of forest was recalculated to % using data on area of Bundeslands
Number of crime cases per capita	Sensch, Jürgen (1955-2003 [2005]), histat-Datenkompilation online: Kriminalitätsentwicklung in der Bundesrepublik Deutschland von 1955 bis 2003: Ausgewählte Indikatoren aus der Kriminalstatistik; Polizeiliche Kriminalstatistik 2005, 2015.	Number of recorded cases	
Number of sever crime cases	Polizeiliche Kriminalstatistik 1974, 1978, 1982, 1986, 1990, 1994, 1998, 2002, 2006, 2010, 2014, 2018	Number of murders	
GDP per capita	Statistisches Landesamt Baden-Württemberg, Arbeitskreis "Volkswirtschaftliche Gesamtrechnungen der Länder", "Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern der Bundesrepublik Deutschland" 1961-2020	All Bundeslands	
Employment per capita	Erwerbstätige in den alten Ländern der Bundesrepublik Deutschland 1970 bis 1991 sowie in deren kreisfreien Städten und Landkreisen 1980, 1985, 1987 bis 1991, Destatis 2005; Erwerbstätige am Arbeitsort Ländererergebnisse – Jahresdurchschnitt, Destatis 1991-2020	All Bundeslands	

	Male			Female		
	Single	Inter- -married	Noninter -married	Single	Inter- -married	Noninter -married
Age	38.08	39.65	39.11	41.97	38.28	36.21
	(8.71)	(9.00)	(8.44)	(10.18)	(8.09)	(7.65)
Years since mig.	16.57	20.14	19.54	18.89	17.15	17.75
	(11.39)	(12.25)	(9.94)	(12.29)	(10.55)	(9.24)
Mig. age	21.41	19.20	19.23	23.22	20.67	18.30
	(12.62)	(10.53)	(8.74)	(11.97)	(10.29)	(8.47)
Education	11.41	11.42	10.68	11.14	11.64	10.46
	(2.26)	(2.52)	(2.33)	(2.65)	(2.56)	(2.19)
Feel German	2.69	2.93	2.95	2.78	2.73	3.08
	(1.28)	(1.21)	(1.29)	(1.31)	(1.25)	(1.33)
Feel of not belong.	3.91	3.75	3.45	3.30	3.81	3.39
	(0.97)	(1.17)	(1.19)	(0.98)	(1.17)	(1.16)
Oral German skills	2.14	1.81	2.17	1.90	1.87	2.19
	(1.00)	(0.92)	(0.89)	(0.95)	(0.89)	(1.03)
Language of Media	3.66	3.73	3.55	3.89	4.15	3.77
	(1.29)	(1.30)	(1.25)	(1.24)	(1.21)	(1.35)
Visit from German	0.79	0.94	0.84	0.89	0.95	0.84
	(0.41)	(0.24)	(0.36)	(0.31)	(0.23)	(0.36)
Share	16.55%	39.82%	43.63%	13.81%	47.53%	38.66%

Note: The table lists the mean for demographic characteristics calculated with sample population weights. Standard deviation in the parentheses. Variable Feel German is measured on a 1-5 scale, where 1 means "Completely" and 5 means "Not at All". Variable Feel of not Belonging is measured on a 1-5 scale, where 1 means "Very often" and 5 means "Never". Variable Oral German skills is measured on a 1-5 scale, where 1 means "Very Good" and 5 means "Not at All". Variable Language of Media is measured on a scale 1-5 where 1 means only language country of origin and 5 means only German. Variable Received Visits of Germans is an indicator variable equal 1 if the agent received at least one visit from a German in the previous year.

Table 15: Descriptive Statistics of Immigrants Subsample

C The solution of single and couple households' problems at the Stage 3

C.1 Couple's problem

This section provides the solution for stage 3 for married couples deciding on their private and public consumption and labor supply.

Individual utilities of agents in couple (H, H^*) at Stage 3 are:

$$u(Q, C, L) = \ln Q(C + \alpha \cdot \ell_{nw}) \tag{26}$$
$$u(Q, C^*, L^*) = \ln Q(C^* + \alpha^* \cdot (\ell^*_{pt} + \ell^*_{nw}) + \delta^* \cdot \ell^*_{pt}) \tag{27}$$

Preferences satisfy the *transferable utility* (TU) if there exists cardinalisation of representing them utilities, such that far all values of prices and income, the Pareto frontier is a straight line with a slope equal to -1 (Chiappori and Gugl, 2020).

Proposition 3. *Preferences represented by the utility given in Equations 26 and 27 satisfy TU property.*

Proof: Assume that we take cardinal representation of the preferences equal to $\exp u_i$ and $\exp u^*$. Then the couple maximisation problem can be written in the following form:

$$\max_{Q,C,C^*} \exp u + \mu \exp u^* \tag{28}$$

under the budget constraint:

$$\overline{Y}^{H,H^*}(L, L^*) \geq C + C^* + pQ \tag{29}$$

where μ is a Pareto weight and Y represents income of a couple (H, H^*) as a function on their leisure choices. Then using Lagrangia function, the problem can be expressed as:

$$\begin{aligned} \mathcal{L}(Q, C, C^*, \lambda) = Q(C + C^* + \alpha \cdot \ell_{nw} + \alpha^* \cdot (\ell^*_{pt} + \ell^*_{nw}) + \delta^* \cdot \ell^*_{pt}) \\ + \lambda(\overline{Y}^{H,H^*}(L, L^*) - C - C^* - pQ) \end{aligned} \tag{30}$$

Taking the derivatives with respect to private consumptions yields:

$$\begin{cases} \frac{\partial L}{\partial C} & = Q - \lambda = 0 \\ \frac{\partial L}{\partial C^*} & = \mu Q - \lambda = 0 \end{cases} \implies \mu = 1 \tag{31}$$

Transferable utility implies that household aggregate demand does not depend on Pareto weight μ. So, the household (H, H^*) maximization problem at Stage 3 is as follows:

$$\max_{\overline{C},Q} Q(\overline{C} + \Lambda_m(L) + \Lambda_f(L^*)) \tag{32}$$

where

$$\Lambda(L)_g \equiv \begin{cases} \alpha \ell_{nw} & \text{if } g = m \\ \alpha \ell_{nw} + \delta \ell_{pt} & \text{otherwise} \end{cases} \tag{33}$$

with respect to the budget constraint:

$$\overline{Y}^{H,H^*}(L,L^*) \equiv y_{nl}(H,H^*) + \ell_{nw}\cdot b(w) + \ell^*_{nw}\cdot b(w^*) + w_{net}(L,L^*,w,w^*) \tag{34}$$
$$= \overline{C} + pQ. \tag{35}$$

Conditioning on labor supply (L, L^*) the ex-post (after realization of productivity and leisure preference shocks) efficient allocation is as follows:

$$pQ(L,L^*) = (\overline{Y}^{H,H^*}(L,L^*) + \Lambda_m(L) + \Lambda_f(L^*))/2 \tag{36}$$

$$\overline{C}(L,L^*) = (\overline{Y}^{H,H^*}(L,L^*) - \Lambda_m(L) - \Lambda_f(L^*))/2 = pQ - \Lambda_m(L) - \Lambda_f(L^*) \tag{37}$$

The equations 36 and 37 describe the aggregated demand of couple (H, H^*) for private and public consumptions as a function of individual labor supply choices. Using this fact, the optimal labor supply can be found by solving the following maximization problem:

$$\max_{L,L^*} pQ^2(L,L^*). \tag{38}$$

The final maximization problem is a discrete choice problem. Each couple (H, H^*) has 3×2 possible labor supply choices. Given the solution to this problem, one can recover aggregated demands for private and public consumptions of union (H, H^*).

Let's define $C^*(L,L^*) = (\overline{Y}^{H,H^*}(L,L^*) - \Lambda_m(L) - \Lambda_f(L^*))/2 - C(L,L^*)$. Ex-ante (at Stage 2, before realization of productivity and leisure preference shocks) efficiency requires that C maximizes some weighted sum of individual expected utilities, formally:

$$\max_{C} \mathrm{E}\, u + \mu\, \mathrm{E}\, u^* \tag{39}$$

for some $\mu > 0$, under the resources constraint given by 34. First-order condition implies:

$$\frac{\partial}{\partial C} = \frac{1}{C+\Lambda_m(L)} - \frac{\mu}{(\overline{Y}^{H,H^*}(L,L^*) - \Lambda_m(L) - \Lambda_f(L^*))/2 - C - \Lambda_f(L^*)} = 0 \tag{40}$$

As a result, private consumption of agents is given by:

$$C = \frac{\overline{Y}^{H,H^*}(L,L^*) + \Lambda_m(L) + \Lambda_f(L^*)}{2(\mu+1)} - \Lambda(L) = \frac{1}{1+\mu} pQ(L,L^*) - \Lambda_m(L) \tag{41}$$

$$C^* = \mu\cdot \frac{\overline{Y}^{H,H^*}(L,L^*) + \Lambda_m(L) + \Lambda_f(L^*)}{2(\mu+1)} - \Lambda(L^*) = \frac{\mu}{1+\mu} pQ(L,L^*) - \Lambda_f(L^*) \tag{42}$$

$$\tag{43}$$

Finally, individual expected utilities are equal to the following:

$$\mathrm{E}\, u = \ln p + \ln \frac{1}{1+\mu} + \int \ln Q^2(H, H^*, r_{ij}, \boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}) dF(\boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}) \tag{44}$$

$$\mathrm{E}\, u^* = \ln p + \ln \frac{\mu}{1+\mu} + \int \ln Q^2(H, H^*, r_{ij}, \boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}) dF(\boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}) \tag{45}$$

where F denotes the joint distribution of productivity and leisure shocks.

Let $\Psi(H, H^*, r)$ denotes the common part of private consumption:

$$\Psi(H, H^*, r) = \ln p + \int \ln Q^2(H, H^*, r_{ij}, \boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}) dF(\boldsymbol{\varepsilon}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}). \tag{46}$$

Then, taking exp of both sides and adding up gives the set of ex-ante (at Stage 2) Pareto efficient allocations:

$$\exp\{\mathrm{E}\, u\} + \exp\{\mathrm{E}\, u^*\} = \frac{1}{1+\mu} \exp\{\Psi(H, H^*, r)\} + \frac{\mu}{1+\mu} \exp\{\Psi(H, H^*, r)\} \tag{47}$$

$$= \exp\{\Psi(H, H^*, r)\} = \overline{U}(H, H^*, r) \tag{48}$$

which is a TU form for $U_g^H(H^*, r) = \exp\{\mathrm{E}\, u\}$ and $U_{g^*}^{H^*}(H, r) = \exp\{\mathrm{E}\, u^*\}$.

C.2 Single's problem

This section provides the solution for stage 3 for singles who choose their private and public consumption and labor supply. Single agents at stage 3 face the following maximization problem:

At the stage 3 of the model single household $(H, \varnothing)$ (equivalently for $(\varnothing, H_j)$) solves the following maximization problem:

$$\max_{C,Q} Q(C + \Lambda_g(L)) \tag{49}$$

with respect to the budget constrain:

$$Y^H(L) \equiv y_{nl}(H) + \ell_{nw} \cdot b(w) + w_{net}(w, L) = C + pQ. \tag{50}$$

where $\Lambda_g(L)$ is defined as in Equation 33.

The maximization problem of single agents is very similar to that of a couple since the union of (H, H^*) at stage 3 behaves as a single decision maker. Then, the conditional on the labor supply choice demand for private and public goods is given by:

$$pQ(L) = (Y^H(L) + \Lambda(L))/2 \tag{51}$$

$$C(L) = (Y^H(L) - \Lambda(L))/2 = pQ - \Lambda(L) \tag{52}$$

Using Equations 51 and 52 to substitute Q and C in the Equation 49, one gets the expression, which can be used to find the optimal labor supply. Single i finds the optimal labor supply by solving the following maximization problem:

$$\max_{L} pQ^2(L) \tag{53}$$

It is a discrete choice problem, where every single agent has three (or two for men) possible choices. Given the solution to this problem, one can recover demand for the single agent's public and private consumption.

Then expected utility (at Stage 2, before realization of productivity and leisure preference shocks) is given by:

$$\mathrm{E}\, u = \ln p + \int \ln Q^2(H, r, \varepsilon, \upsilon, \zeta) dF(\varepsilon, \upsilon, \zeta)$$

Finally define the exponential representation of the utility function $U_g^H(\varnothing, r) \equiv \exp\{\mathrm{E}\, u\}$, which corresponds to a TU form from the marriage problem.

D Identification of marriage market parameters and sharing rule

Let N_r^H be a number of men with human capital H who live in region r. Then, $N_{d,r}^{H,H^*}$ is the number of (H, H^*) marriages demanded by men with human capital H and $N_{d,r}^{H,\varnothing}$ is the number of unmarried men with human capital H. Using Equation 20 and the fact that ML estimator of $\mathrm{P}(H^*|H, r)$ is $\frac{N_{d,r}^{H,H^*}}{N_r^H}$, I derive a quasi-demand equation for men:

$$\begin{aligned} \ln \frac{N_{d,r}^{H,H^*}}{N_{d,r}^{H,\varnothing}} &= \Gamma_M(H, H^, r*) - \Gamma_M(H, \varnothing, r) = \\ &= (\overline{U}(H, H^*, r) - \tau(H, H^*, r) - U_g(H, \varnothing, r) + \phi_1|e^* - e| + \phi_2|o^* - o| - \phi_{0H})/\sigma_\omega^{M,H} \end{aligned} \tag{54}$$

and a quasi-supply equation for women:

$$\begin{aligned} \ln \frac{N_{s,r}^{H,H^*}}{N_{s,r}^{\varnothing,H^*}} &= \Gamma_F(H^*, H, r) - \Gamma_F(H^*, \varnothing, r) \\ &= (\tau(H, H^*, r) - U_{g^*}(H^*, \varnothing, r) + \phi_1|e - e^*| + \phi_2|o - o^*| - \phi_{0H^*})/\sigma_\omega^{F,H^*} \end{aligned} \tag{55}$$

In each location, r 4×4 sub-marriage market clears, when given equilibrium transfers τ's, the demand by men with H for women with H^* is equal to the supply of women with H^* for men with H for all possible combinations of the human capital. Finally, the identification of transfers τ's and taste for similarity *phi*'s can be obtained using Equations 54 and 55.

To identify μ's so Pareto weights associated with the initial maximization problem, I use Equation 47 and show that:

$$\tau(H, H^*, r) = U_{g^*}(H^*, H, r) = \frac{\mu(H, H^*, r)}{1 + \mu(H, H^*, r)} \times \overline{U}(H, H^*, r) \tag{56}$$

Solving for $\mu(H, H^*, r)$ yields:

$$\mu(H, H^*, r) = \frac{\tau(H, H^*, r)}{\overline{U}(H, H^*, r) - \tau(H, H^*, r)} \tag{57}$$

$\mu(H, H^*, r)$ is well-defined if $\tau(H, H^*, r) \in (0, \overline{U}(H, H^*, r))$. So, if $\tau(H, H^*, r)$ is identified, then the $\mu(H, H^*, r)$ is also identified.

E German social security and tax systems

E.1 German tax code

In Germany, each employee pays two types of social contribution: social system contribution and personal income tax. Social system contribution depends on individual yearly labor income. However, married couples in Germany submit tax statements together. As a result, the amount of paid personal income tax depends on the yearly labor income of both partners.

To approximate the level of social contribution, first, I define yearly individual labor income as a function of wage and labor supply choices. I assume that every full-time employed agent works approximately 1778 hours per year, while the part-time employed agent works half of it. So, the individual yearly labor income y is obtained in the following way:

$$y(w, L) = w \cdot [(\ell_{ft} + 0.5\ell_{pt}) \cdot 1778]$$

In Germany, individuals who earn less than 4.800€ do not pay social system contributions. There is also a maximum amount to contribute to the social system. This amount slightly changes every year. Since this paper uses data from 1984 to 2018, I take the threshold from 2005 (13104€) as a representative for the whole sample.

Then, I use the following piece-wise function of yearly individual labor income to approximate the share of gross income which contributes to the social system:

$$\tau_{sc}(y) = \begin{cases} 0 & \text{if } y \leq 4800 \\ 0.0002625 \cdot y & \text{if } 4800 < y \leq 9600 \\ 0.21 & \text{if } 9600 < y \leq 62400 \\ 13104/y & \text{otherwise} \end{cases}$$

The income tax rate is calculated using taxable income, which I take as income after the social security contribution. For agent with H it is $y^H_{sc} = (1 - \tau_{sc}(y^H)) \cdot y^H$. As I mentioned, married couples are taxed jointly in Germany as if each earned half of the joint income. This situation can be especially beneficial when there is a big gap between partners' incomes. It provides incentives for one of the partners to work less.

Assume that y_{sc}^{H,H^*} is an average taxable couple's yearly labor income. The tax schedule changes slightly in Germany every year. For consistency, I use a tax schedule for 2005 as a representative for my sample. So, the tax rate of individuals in a couple (H, H^*) is approximated in the following way:

$$\tau_{pit}(y_{sc}^{H,H^*}) = \begin{cases} 0 & \text{if } y_{sc}^{H,H^*} \leq 7664 \\ ((883.74 \cdot \hat{y}^{H,H^*} + 1500) \cdot \hat{y}^{H,H^*})/y_{sc}^{H,H^*} & \text{if } 7664 < y_{sc}^{H,H^*} \leq 12740 \\ ((228.74 \cdot \bar{y}^{H,H^*} + 2397) \cdot \bar{y}^{H,H^*} + 989)/y_{sc}^{H,H^*} & \text{if } 12740 < y_{sc}^{H,H^*} < 52152 \\ (0.42 \cdot y_{sc}^{H,H^*} - 7914)/y_{sc}^{H,H^*} & \text{otherwise} \end{cases}$$

where:

$$\hat{y}^{H,H^*} = (y_{sc}^{H,H^*} - 7644)/10000 \qquad \bar{y}^{H,H^*} = (y_{sc}^{H,H^*} - 12740)/10000$$

In case of single agents y_{sc}^{H,H^*} is replaced by individual taxable yearly labor income y_{sc}^{H}.

E.2 Unemployment benefit

In Germany, unemployment benefit is a percentage of the last obtained income and is bounded from above by a certain threshold set by the government. The rules determining the size of unemployment benefit change over time. Since in this paper I use data from 1984 to 2017, I take the threshold (8.68€/h) and percentage of the last obtained income (60%) from 2005 as a representative for the whole sample.

Given the static structure of the model, there is no information about the last period's income. To approximate an unemployment benefit, first, I approximate the last obtained income using an expected wage $\mathrm{E}[w]$ (so wage w net of productivity shock) and assume that an agent worked full-time. Then, I calculate the size of an unemployment benefit using the following formula:

$$b(w) = \min(0.6 \cdot w_{net}(\ell_{ft} = 1, E[w]), 8.68) \tag{58}$$

F Model fit

This subsection of appendix contains a set of tables showing the fitness of the model. Header "Simulation" refers to moments obtained in the simulation. The data moments are included under header "Data", while their standard errors are presented under the header "Data SE". Finally, header "Diff in SE" corresponds to the difference between simulated and data moments expressed in standard deviations.

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.695	2.693	0.004	0.548
Variance	0.158	0.208	0.052	0.951
P(wage<Q10)	0.101	0.100	0.009	0.103
P(wage<Q25)	0.307	0.250	0.005	11.864
P(wage<Q50)	0.538	0.500	0.004	9.644
P(wage<Q75)	0.748	0.750	0.005	0.403
P(wage<Q90)	0.888	0.900	0.006	1.885
College				
Mean	3.202	3.197	0.008	0.612
Variance	0.246	0.229	0.072	0.229
P(wage<Q10)	0.142	0.100	0.019	2.209
P(wage<Q25)	0.355	0.250	0.008	12.359
P(wage<Q50)	0.540	0.500	0.007	5.468
P(wage<Q75)	0.715	0.750	0.009	4.053
P(wage<Q90)	0.855	0.900	0.013	3.517

Table 17: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.600	2.583	0.012	1.314
Variance	0.142	0.210	0.089	0.756
P(wage<Q10)	0.101	0.100	0.022	0.043
P(wage<Q25)	0.300	0.250	0.014	3.629
P(wage<Q50)	0.540	0.500	0.011	3.590
P(wage<Q75)	0.741	0.750	0.013	0.724
P(wage<Q90)	0.879	0.900	0.016	1.268
College				
Mean	3.136	3.112	0.035	0.667
Variance	0.194	0.291	0.165	0.588
P(wage<Q10)	0.046	0.100	0.075	0.715
P(wage<Q25)	0.276	0.255	0.052	0.419
P(wage<Q50)	0.585	0.500	0.038	2.233
P(wage<Q75)	0.773	0.750	0.038	0.591
P(wage<Q90)	0.891	0.905	0.045	0.311

Table 18: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.383	2.378	0.006	0.883
Variance	0.271	0.357	0.070	1.225
P(wage<Q10)	0.072	0.100	0.015	1.850
P(wage<Q25)	0.254	0.250	0.008	0.444
P(wage<Q50)	0.547	0.500	0.006	7.588
P(wage<Q75)	0.763	0.750	0.005	2.468
P(wage<Q90)	0.887	0.900	0.007	1.794
College				
Mean	2.939	2.940	0.013	0.057
Variance	0.210	0.353	0.105	1.367
P(wage<Q10)	0.060	0.100	0.035	1.160
P(wage<Q25)	0.290	0.250	0.017	2.390
P(wage<Q50)	0.563	0.500	0.012	5.015
P(wage<Q75)	0.786	0.750	0.013	2.769
P(wage<Q90)	0.910	0.900	0.018	0.541

Table 19: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.302	2.289	0.017	0.741
Variance	0.293	0.315	0.117	0.188
P(wage<Q10)	0.115	0.100	0.037	0.410
P(wage<Q25)	0.279	0.250	0.021	1.347
P(wage<Q50)	0.480	0.500	0.019	1.101
P(wage<Q75)	0.731	0.751	0.020	0.963
P(wage<Q90)	0.871	0.900	0.023	1.265
College				
Mean	2.763	2.745	0.039	0.444
Variance	0.239	0.446	0.193	1.073
P(wage<Q10)	0.037	0.101	0.084	0.759
P(wage<Q25)	0.230	0.251	0.057	0.373
P(wage<Q50)	0.519	0.502	0.047	0.371
P(wage<Q75)	0.809	0.750	0.045	1.306
P(wage<Q90)	0.946	0.900	0.050	0.921

Table 20: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.604	2.553	0.015	3.374
Variance	0.138	0.281	0.106	1.350
P(wage<Q10)	0.036	0.100	0.035	1.858
P(wage<Q25)	0.223	0.250	0.019	1.432
P(wage<Q50)	0.550	0.500	0.015	3.298
P(wage<Q75)	0.782	0.750	0.014	2.195
P(wage<Q90)	0.898	0.900	0.016	0.153
College				
Mean	3.033	3.001	0.024	1.329
Variance	0.224	0.275	0.133	0.385
P(wage<Q10)	0.062	0.100	0.049	0.788
P(wage<Q25)	0.297	0.251	0.031	1.475
P(wage<Q50)	0.548	0.502	0.026	1.785
P(wage<Q75)	0.742	0.750	0.027	0.284
P(wage<Q90)	0.877	0.901	0.036	0.644

Table 21: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.406	2.406	0.048	0.006
Variance	0.161	0.222	0.178	0.344
P(wage<Q10)	0.087	0.105	0.093	0.191
P(wage<Q25)	0.241	0.251	0.062	0.159
P(wage<Q50)	0.485	0.501	0.055	0.279
P(wage<Q75)	0.768	0.755	0.056	0.222
P(wage<Q90)	0.916	0.908	0.067	0.120
College				
Mean	3.115	3.078	0.090	0.414
Variance	0.182	0.324	0.271	0.526
P(wage<Q10)	0.073	0.103	0.109	0.278
P(wage<Q25)	0.265	0.264	0.104	0.010
P(wage<Q50)	0.557	0.502	0.113	0.490
P(wage<Q75)	0.796	0.756	0.107	0.374
P(wage<Q90)	0.902	0.918	0.099	0.162

Table 22: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.495	2.461	0.015	2.297
Variance	0.241	0.325	0.109	0.766
P(wage<Q10)	0.078	0.100	0.041	0.540
P(wage<Q25)	0.298	0.250	0.020	2.347
P(wage<Q50)	0.562	0.500	0.013	4.656
P(wage<Q75)	0.731	0.751	0.013	1.586
P(wage<Q90)	0.850	0.900	0.016	3.253
College				
Mean	2.932	2.891	0.022	1.880
Variance	0.218	0.301	0.131	0.633
P(wage<Q10)	0.062	0.100	0.064	0.605
P(wage<Q25)	0.308	0.250	0.029	2.020
P(wage<Q50)	0.564	0.500	0.020	3.178
P(wage<Q75)	0.724	0.750	0.020	1.305
P(wage<Q90)	0.866	0.901	0.027	1.281

Table 23: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.285	2.299	0.052	0.269
Variance	0.231	0.360	0.210	0.614
P(wage<Q10)	0.093	0.102	0.072	0.125
P(wage<Q25)	0.221	0.260	0.064	0.610
P(wage<Q50)	0.507	0.508	0.063	0.020
P(wage<Q75)	0.784	0.751	0.058	0.576
P(wage<Q90)	0.882	0.901	0.053	0.360
College				
Mean	2.750	2.732	0.082	0.221
Variance	0.202	0.360	0.266	0.594
P(wage<Q10)	0.039	0.105	0.187	0.352
P(wage<Q25)	0.199	0.267	0.115	0.586
P(wage<Q50)	0.544	0.510	0.109	0.308
P(wage<Q75)	0.840	0.755	0.098	0.860
P(wage<Q90)	0.896	0.903	0.079	0.083

Table 24: Log wage, married, immigrant male

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.695	2.691	0.004	0.830
		Variance	0.158	0.208	0.053	0.950
	immig.	Mean	2.700	2.716	0.017	0.905
		Variance	0.159	0.198	0.103	0.380
College	native	Mean	3.198	3.195	0.008	0.340
		Variance	0.245	0.229	0.073	0.222
	immig.	Mean	3.237	3.215	0.029	0.740
		Variance	0.251	0.234	0.142	0.119

Table 25: Log wage, native male

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.614	2.591	0.021	1.097
		Variance	0.141	0.257	0.123	0.946
	immig.	Mean	2.588	2.577	0.014	0.784
		Variance	0.143	0.167	0.090	0.265
College	native	Mean	3.158	3.132	0.041	0.611
		Variance	0.189	0.273	0.175	0.479
	immig.	Mean	3.112	3.087	0.062	0.414
		Variance	0.198	0.314	0.222	0.526

Table 26: Log wage, immigrant male

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.382	2.376	0.006	0.876
		Variance	0.272	0.357	0.071	1.206
	immig.	Mean	2.405	2.403	0.027	0.073
		Variance	0.263	0.354	0.150	0.608
College	native	Mean	2.937	2.938	0.013	0.103
		Variance	0.210	0.348	0.106	1.306
	immig.	Mean	2.963	2.958	0.055	0.081
		Variance	0.209	0.418	0.225	0.929

Table 27: Log wage, native female

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.312	2.290	0.027	0.827
		Variance	0.298	0.334	0.148	0.244
	immig.	Mean	2.291	2.288	0.020	0.168
		Variance	0.287	0.292	0.125	0.037
College	native	Mean	2.743	2.725	0.047	0.386
		Variance	0.243	0.478	0.215	1.093
	immig.	Mean	2.813	2.804	0.070	0.129
		Variance	0.224	0.354	0.244	0.533

Table 28: Log wage, immigrant female

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	2.635	2.614	0.009	2.376
North	College	3.102	3.085	0.016	1.088
South	Noncollege	2.711	2.702	0.006	1.574
South	College	3.202	3.204	0.011	0.142
West	Noncollege	2.683	2.684	0.007	0.101
West	College	3.188	3.166	0.013	1.722

Table 29: Mean of log wage by region and education,native male

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	2.224	2.184	0.042	0.956
North	College	2.565	2.529	0.105	0.350
South	Noncollege	2.358	2.339	0.023	0.812
South	College	2.852	2.807	0.048	0.937
West	Noncollege	2.257	2.265	0.028	0.291
West	College	2.713	2.755	0.056	0.755

Table 30: Mean of log wage by region and education,immigrant male

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	0.868	0.869	0.003	0.280
North	College	0.895	0.896	0.005	0.215
South	Noncollege	0.715	0.714	0.004	0.262
South	College	0.831	0.857	0.007	3.566
West	Noncollege	0.711	0.711	0.011	0.006
West	College	0.832	0.826	0.015	0.349

Table 31: Mean of log wage by region and education,native female

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	0.649	0.636	0.032	0.404
North	College	0.797	0.756	0.050	0.827
South	Noncollege	0.388	0.319	0.011	6.156
South	College	0.375	0.344	0.023	1.311
West	Noncollege	0.178	0.223	0.028	1.610
West	College	0.153	0.229	0.049	1.560

Table 32: Mean of log wage by region and education,immigrant female

		Simulation	Data	Data SE	Diff in SE
Male					
	native x noncollege	0.868	0.869	0.003	0.280
	native x college	0.895	0.896	0.005	0.215
	immig x noncollege	0.730	0.747	0.011	1.564
	immig x college	0.681	0.693	0.026	0.428
Female					
	native x noncollege	0.715	0.714	0.004	0.262
	native x college	0.831	0.857	0.007	3.566
	immig x noncollege	0.584	0.578	0.012	0.496
	immig x college	0.640	0.639	0.023	0.036

Table 33: Probability of working ($\ell^{ft} + \ell^{pt} = 1$), married agents

		Simulation	Data	Data SE	Diff in SE
Male					
	native x noncollege	0.711	0.711	0.011	0.006
	native x college	0.832	0.826	0.015	0.349
	immig x noncollege	0.622	0.571	0.037	1.373
	immig x college	0.633	0.617	0.061	0.264
Female					
	native x noncollege	0.827	0.827	0.009	0.038
	native x college	0.931	0.936	0.009	0.481
	immig x noncollege	0.649	0.636	0.032	0.404
	immig x college	0.797	0.756	0.050	0.827

Table 34: Probability of working ($\ell^{ft} + \ell^{pt} = 1$), single agents

Region	Education	Simulation	Data	Data SE	Diff in SE
Married					
	native x noncollege	0.380	0.378	0.004	0.516
	native x college	0.275	0.338	0.010	6.655
	immig x noncollege	0.388	0.319	0.011	6.156
	immig x college	0.375	0.344	0.023	1.311
Single					
	native x noncollege	0.164	0.175	0.009	1.285
	native x college	0.134	0.150	0.014	1.160
	immig x noncollege	0.178	0.223	0.028	1.610
	immig x college	0.153	0.229	0.049	1.560

Table 35: Probability of part-time working ($\ell^{pt} = 1$), females

	Female Human Capital				
Male Human Capital	Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.761	0.040	0.030	0.004	0.165
Native coll.	0.552	0.282	0.011	0.016	0.139
Immig. noncoll.	0.314	0.026	0.461	0.038	0.161
Immig. coll.	0.154	0.141	0.134	0.295	0.275
Data					
Native noncoll.	0.749	0.043	0.031	0.004	0.174
Native coll.	0.428	0.366	0.016	0.022	0.168
Immig. noncoll.	0.293	0.024	0.454	0.079	0.150
Immig. coll.	0.160	0.201	0.126	0.364	0.149
Diff in SE					
Native noncoll.	1.681	0.877	0.155	0.092	1.370
Native coll.	8.890	6.233	1.363	1.428	2.713
Immig. noncoll.	0.990	0.280	0.271	3.181	0.648
Immig. coll.	0.183	1.638	0.276	1.592	3.932

Table 36: Probability of marriage, male born in ’50s living in North

Male Human Capital	Female Human Capital				
	Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.780	0.042	0.036	0.004	0.138
Native coll.	0.588	0.335	0.017	0.021	0.038
Immig. noncoll.	0.296	0.024	0.526	0.035	0.118
Immig. coll.	0.175	0.169	0.197	0.417	0.041
Data					
Native noncoll.	0.800	0.034	0.039	0.004	0.123
Native coll.	0.515	0.319	0.030	0.027	0.109
Immig. noncoll.	0.273	0.022	0.562	0.057	0.086
Immig. coll.	0.179	0.159	0.192	0.370	0.101
Diff in SE					
Native noncoll.	4.286	3.647	1.267	0.252	4.000
Native coll.	7.252	1.739	3.710	2.046	11.137
Immig. noncoll.	1.951	0.615	2.755	3.583	4.429
Immig. coll.	0.158	0.484	0.240	1.688	3.412

Table 37: Probability of marriage, male born in '50s living in South

Male Human Capital	Female Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.745	0.041	0.033	0.004	0.178
Native coll.	0.541	0.306	0.014	0.017	0.122
Immig. noncoll.	0.284	0.024	0.471	0.033	0.188
Immig. coll.	0.158	0.160	0.161	0.311	0.212
Data					
Native noncoll.	0.807	0.033	0.033	0.004	0.124
Native coll.	0.523	0.332	0.021	0.020	0.105
Immig. noncoll.	0.283	0.022	0.562	0.052	0.081
Immig. coll.	0.179	0.161	0.180	0.385	0.095
Diff in SE					
Native noncoll.	11.959	3.501	0.042	0.099	12.489
Native coll.	1.518	2.238	1.935	0.859	2.259
Immig. noncoll.	0.064	0.419	5.707	2.675	12.214
Immig. coll.	0.726	0.037	0.648	1.990	5.140

Table 38: Probability of marriage, male born in '50s living in West

Male Human Capital	Female Human Capital				
	Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.686	0.045	0.031	0.005	0.234
Native coll.	0.513	0.267	0.014	0.016	0.190
Immig. noncoll.	0.274	0.030	0.454	0.042	0.199
Immig. coll.	0.157	0.136	0.145	0.278	0.283
Data					
Native noncoll.	0.666	0.047	0.027	0.006	0.255
Native coll.	0.392	0.346	0.019	0.029	0.214
Immig. noncoll.	0.242	0.030	0.480	0.073	0.176
Immig. coll.	0.155	0.138	0.146	0.319	0.243
Diff in SE					
Native noncoll.	3.189	0.974	2.116	1.002	3.590
Native coll.	10.554	7.018	1.555	3.302	2.554
Immig. noncoll.	1.994	0.006	1.350	3.096	1.642
Immig. coll.	0.086	0.057	0.025	1.124	1.215

Table 39: Probability of marriage, male born in '60s living in North

Male Human Capital	Female Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.712	0.047	0.038	0.005	0.199
Native coll.	0.568	0.330	0.023	0.021	0.058
Immig. noncoll.	0.260	0.029	0.522	0.040	0.149
Immig. coll.	0.179	0.163	0.214	0.402	0.042
Data					
Native noncoll.	0.733	0.042	0.039	0.006	0.180
Native coll.	0.469	0.309	0.028	0.033	0.160
Immig. noncoll.	0.235	0.022	0.579	0.053	0.111
Immig. coll.	0.200	0.134	0.205	0.320	0.140
Diff in SE					
Native noncoll.	4.832	2.567	0.785	1.764	4.970
Native coll.	12.403	2.839	2.187	4.137	17.439
Immig. noncoll.	2.656	1.956	5.114	2.630	5.426
Immig. coll.	1.128	1.789	0.433	3.674	5.907

Table 40: Probability of marriage, male born in '60s living in South

Male Human Capital	Female Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.668	0.045	0.033	0.004	0.249
Native coll.	0.505	0.293	0.017	0.017	0.168
Immig. noncoll.	0.248	0.028	0.459	0.037	0.228
Immig. coll.	0.160	0.154	0.170	0.298	0.219
Data					
Native noncoll.	0.736	0.043	0.035	0.004	0.183
Native coll.	0.489	0.319	0.021	0.022	0.149
Immig. noncoll.	0.232	0.021	0.572	0.049	0.126
Immig. coll.	0.202	0.180	0.198	0.306	0.114
Diff in SE					
Native noncoll.	13.364	1.191	0.500	0.162	14.833
Native coll.	1.569	2.683	1.188	1.795	2.525
Immig. noncoll.	1.432	1.815	8.456	2.092	11.363
Immig. coll.	1.619	1.050	1.055	0.265	5.028

Table 41: Probability of marriage, male born in '60s living in West

Male Human Capital	Female Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.616	0.054	0.057	0.008	0.265
Native coll.	0.382	0.291	0.026	0.033	0.267
Immig. noncoll.	0.223	0.028	0.487	0.048	0.214
Immig. coll.	0.099	0.129	0.141	0.283	0.348
Data					
Native noncoll.	0.582	0.061	0.049	0.010	0.298
Native coll.	0.307	0.357	0.023	0.050	0.264
Immig. noncoll.	0.196	0.025	0.532	0.057	0.190
Immig. coll.	0.094	0.133	0.156	0.332	0.286
Diff in SE					
Native noncoll.	4.215	1.969	2.234	1.148	4.319
Native coll.	6.242	5.165	0.724	2.917	0.280
Immig. noncoll.	1.956	0.686	2.615	1.080	1.711
Immig. coll.	0.278	0.164	0.571	1.451	1.915

Table 42: Probability of marriage, male born in ’70s living in North

Male Human Capital	Female Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.640	0.057	0.069	0.008	0.227
Native coll.	0.440	0.376	0.042	0.049	0.093
Immig. noncoll.	0.211	0.027	0.558	0.046	0.158
Immig. coll.	0.119	0.163	0.218	0.443	0.058
Data					
Native noncoll.	0.626	0.064	0.074	0.013	0.222
Native coll.	0.337	0.341	0.042	0.058	0.222
Immig. noncoll.	0.243	0.023	0.551	0.056	0.128
Immig. coll.	0.113	0.134	0.185	0.361	0.206
Diff in SE					
Native noncoll.	2.211	2.092	1.742	3.794	0.805
Native coll.	11.914	4.078	0.096	2.252	16.981
Immig. noncoll.	3.358	1.442	0.603	1.884	4.064
Immig. coll.	0.397	1.824	1.885	3.787	8.150

Table 43: Probability of marriage, male born in '70s living in South

Male Human Capital	Female Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.598	0.055	0.060	0.008	0.280
Native coll.	0.375	0.322	0.031	0.036	0.235
Immig. noncoll.	0.203	0.027	0.487	0.043	0.241
Immig. coll.	0.102	0.147	0.165	0.311	0.276
Data					
Native noncoll.	0.627	0.062	0.067	0.010	0.233
Native coll.	0.358	0.362	0.034	0.043	0.203
Immig. noncoll.	0.217	0.016	0.620	0.033	0.114
Immig. coll.	0.154	0.131	0.252	0.319	0.144
Diff in SE					
Native noncoll.	4.213	2.102	1.968	1.631	7.564
Native coll.	1.501	3.386	0.796	1.380	3.315
Immig. noncoll.	1.387	3.220	10.802	2.194	15.785
Immig. coll.	2.308	0.749	3.181	0.277	5.952

Table 44: Probability of marriage, male born in '70s living in West

	Male Human Capital				
Female Human Capital	Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.718	0.150	0.035	0.005	0.092
Native coll.	0.238	0.483	0.018	0.026	0.234
Immig. noncoll.	0.314	0.034	0.562	0.043	0.048
Immig. coll.	0.149	0.182	0.179	0.364	0.127
Data					
Native noncoll.	0.735	0.131	0.030	0.005	0.099
Native coll.	0.201	0.539	0.012	0.032	0.216
Immig. noncoll.	0.355	0.058	0.374	0.123	0.090
Immig. coll.	0.135	0.240	0.036	0.435	0.153
Diff in SE					
Native noncoll.	2.339	3.452	1.951	0.670	1.397
Native coll.	2.753	3.293	1.800	0.936	1.240
Immig. noncoll.	1.772	2.152	7.962	4.985	3.026
Immig. coll.	0.407	1.383	7.738	1.457	0.750

Table 45: Probability of marriage, female born in '50s living in North

	Male Human Capital				
Female Human Capital	Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.714	0.170	0.043	0.007	0.066
Native coll.	0.224	0.568	0.021	0.038	0.148
Immig. noncoll.	0.264	0.040	0.608	0.061	0.028
Immig. coll.	0.105	0.189	0.162	0.510	0.034
Data					
Native noncoll.	0.729	0.154	0.041	0.007	0.069
Native coll.	0.194	0.595	0.021	0.037	0.153
Immig. noncoll.	0.276	0.070	0.492	0.100	0.063
Immig. coll.	0.115	0.272	0.061	0.462	0.089
Diff in SE					
Native noncoll.	3.028	3.949	0.835	0.024	0.974
Native coll.	2.833	1.964	0.022	0.187	0.537
Immig. noncoll.	1.005	4.349	8.563	4.818	5.278
Immig. coll.	0.474	2.862	6.423	1.458	2.946

Table 46: Probability of marriage, female born in '50s living in South

Female Human Capital	Male Human Capital Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.707	0.163	0.037	0.005	0.088
Native coll.	0.226	0.539	0.018	0.029	0.188
Immig. noncoll.	0.293	0.039	0.573	0.047	0.048
Immig. coll.	0.136	0.207	0.175	0.395	0.086
Data					
Native noncoll.	0.751	0.138	0.037	0.005	0.069
Native coll.	0.202	0.587	0.019	0.032	0.161
Immig. noncoll.	0.289	0.051	0.489	0.101	0.069
Immig. coll.	0.142	0.223	0.048	0.486	0.100
Diff in SE					
Native noncoll.	7.987	5.707	0.230	0.394	5.796
Native coll.	1.762	2.965	0.113	0.449	2.303
Immig. noncoll.	0.201	1.624	5.004	5.382	2.404
Immig. coll.	0.204	0.459	6.815	2.084	0.525

Table 47: Probability of marriage, female born in ’50s living in West

Female Human Capital	Male Human Capital Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.665	0.158	0.035	0.005	0.138
Native coll.	0.234	0.444	0.021	0.024	0.279
Immig. noncoll.	0.292	0.043	0.554	0.045	0.066
Immig. coll.	0.159	0.171	0.192	0.320	0.158
Data					
Native noncoll.	0.677	0.125	0.030	0.005	0.163
Native coll.	0.213	0.491	0.016	0.021	0.259
Immig. noncoll.	0.281	0.064	0.442	0.102	0.111
Immig. coll.	0.171	0.275	0.048	0.331	0.175
Diff in SE					
Native noncoll.	1.872	7.339	2.168	0.286	5.144
Native coll.	1.774	3.338	1.169	0.662	1.596
Immig. noncoll.	0.669	2.249	5.766	4.854	3.642
Immig. coll.	0.421	3.036	8.741	0.303	0.566

Table 48: Probability of marriage, female born in '60s living in North

Female Human Capital	Male Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.667	0.183	0.043	0.007	0.099
Native coll.	0.223	0.539	0.024	0.034	0.179
Immig. noncoll.	0.246	0.051	0.602	0.062	0.039
Immig. coll.	0.118	0.190	0.182	0.465	0.045
Data					
Native noncoll.	0.679	0.165	0.038	0.009	0.109
Native coll.	0.200	0.561	0.018	0.030	0.191
Immig. noncoll.	0.254	0.070	0.518	0.094	0.064
Immig. coll.	0.147	0.314	0.067	0.372	0.100
Diff in SE					
Native noncoll.	2.691	5.199	2.648	1.319	3.393
Native coll.	2.675	2.023	1.919	1.309	1.392
Immig. noncoll.	0.760	3.276	7.384	4.854	4.525
Immig. coll.	1.630	5.282	9.115	3.800	3.612

Table 49: Probability of marriage, female born in '60s living in South

Female Human Capital	Male Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.654	0.173	0.036	0.005	0.131
Native coll.	0.222	0.503	0.020	0.026	0.228
Immig. noncoll.	0.274	0.050	0.561	0.049	0.067
Immig. coll.	0.147	0.199	0.189	0.355	0.111
Data					
Native noncoll.	0.706	0.145	0.035	0.006	0.108
Native coll.	0.233	0.538	0.018	0.032	0.179
Immig. noncoll.	0.267	0.050	0.516	0.090	0.077
Immig. coll.	0.168	0.276	0.064	0.398	0.094
Diff in SE					
Native noncoll.	10.261	7.104	0.621	0.877	6.868
Native coll.	0.955	2.698	0.687	1.153	4.856
Immig. noncoll.	0.568	0.013	3.189	5.159	1.370
Immig. coll.	0.760	2.287	6.728	1.165	0.771

Table 50: Probability of marriage, female born in '60s living in West

Female Human Capital	Male Human Capital: Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.635	0.147	0.055	0.006	0.157
Native coll.	0.199	0.404	0.025	0.030	0.342
Immig. noncoll.	0.272	0.046	0.554	0.042	0.087
Immig. coll.	0.125	0.193	0.179	0.273	0.230
Data					
Native noncoll.	0.613	0.121	0.047	0.006	0.213
Native coll.	0.208	0.451	0.019	0.029	0.292
Immig. noncoll.	0.253	0.044	0.503	0.075	0.125
Immig. coll.	0.142	0.269	0.062	0.315	0.212
Diff in SE					
Native noncoll.	2.630	4.887	2.348	0.105	8.204
Native coll.	0.744	3.241	1.561	0.010	3.740
Immig. noncoll.	1.248	0.246	2.914	3.606	3.332
Immig. coll.	0.753	2.610	7.362	1.351	0.671

Table 51: Probability of marriage, female born in '70s living in North

Female Human Capital	Male Human Capital Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.635	0.175	0.067	0.010	0.113
Native coll.	0.193	0.509	0.030	0.045	0.224
Immig. noncoll.	0.231	0.057	0.601	0.060	0.051
Immig. coll.	0.095	0.231	0.175	0.429	0.070
Data					
Native noncoll.	0.618	0.152	0.070	0.009	0.151
Native coll.	0.208	0.509	0.022	0.038	0.224
Immig. noncoll.	0.277	0.071	0.483	0.084	0.085
Immig. coll.	0.151	0.298	0.060	0.345	0.146
Diff in SE					
Native noncoll.	2.843	5.450	0.808	0.136	8.776
Native coll.	1.695	0.020	2.495	1.754	0.042
Immig. noncoll.	4.712	2.525	10.826	3.956	5.575
Immig. coll.	3.646	3.425	11.310	4.147	5.052

Table 52: Probability of marriage, female born in '70s living in South

Female Human Capital	Male Human Capital Native noncoll.	Native coll.	Immig. noncoll.	Immig. coll.	Single
Simulation					
Native noncoll.	0.624	0.162	0.057	0.007	0.151
Native coll.	0.191	0.464	0.025	0.034	0.285
Immig. noncoll.	0.255	0.054	0.557	0.046	0.088
Immig. coll.	0.116	0.230	0.176	0.313	0.166
Data					
Native noncoll.	0.647	0.128	0.071	0.009	0.144
Native coll.	0.242	0.490	0.020	0.030	0.217
Immig. noncoll.	0.235	0.042	0.583	0.068	0.072
Immig. coll.	0.177	0.268	0.078	0.337	0.140
Diff in SE					
Native noncoll.	3.404	6.867	3.819	1.744	1.368
Native coll.	4.218	1.844	1.203	0.725	5.907
Immig. noncoll.	1.858	2.425	2.132	3.446	2.498
Immig. coll.	2.520	1.369	5.776	0.821	1.185

Table 53: Probability of marriage, female born in ’70s living in West

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncoll.	0.244	0.241	0.003	1.197
Native coll.	0.227	0.244	0.006	2.793
Immig. noncoll.	0.205	0.177	0.007	3.845
Immig. coll.	0.211	0.227	0.017	0.902
South				
Native noncoll.	0.407	0.410	0.004	0.877
Native coll.	0.415	0.435	0.007	3.040
Immig. noncoll.	0.456	0.480	0.009	2.560
Immig. coll.	0.473	0.468	0.021	0.239
West				
Native noncoll.	0.349	0.349	0.004	0.168
Native coll.	0.358	0.321	0.006	5.796
Immig. noncoll.	0.339	0.343	0.009	0.397
Immig. coll.	0.316	0.305	0.019	0.562

Table 54: Probability of region choice, male born in '50s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncoll.	0.246	0.251	0.003	1.894
Native coll.	0.229	0.233	0.005	0.702
Immig. noncoll.	0.206	0.195	0.006	1.754
Immig. coll.	0.212	0.221	0.014	0.625
South				
Native noncoll.	0.411	0.412	0.003	0.323
Native coll.	0.417	0.461	0.006	7.932
Immig. noncoll.	0.465	0.465	0.008	0.023
Immig. coll.	0.476	0.497	0.017	1.236
West				
Native noncoll.	0.343	0.336	0.003	2.076
Native coll.	0.353	0.306	0.005	9.226
Immig. noncoll.	0.329	0.340	0.007	1.492
Immig. coll.	0.312	0.282	0.016	1.950

Table 55: Probability of region choice, male born in '60s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncoll.	0.249	0.262	0.004	3.656
Native coll.	0.233	0.245	0.006	2.260
Immig. noncoll.	0.210	0.209	0.006	0.126
Immig. coll.	0.215	0.240	0.014	1.793
South				
Native noncoll.	0.412	0.407	0.004	1.161
Native coll.	0.415	0.466	0.006	7.934
Immig. noncoll.	0.467	0.419	0.007	6.493
Immig. coll.	0.473	0.486	0.016	0.799
West				
Native noncoll.	0.339	0.331	0.004	2.203
Native coll.	0.352	0.289	0.006	10.878
Immig. noncoll.	0.323	0.372	0.007	6.733
Immig. coll.	0.312	0.274	0.015	2.611

Table 56: Probability of region choice, male born in '70s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncoll.	0.242	0.229	0.003	4.100
Native coll.	0.228	0.285	0.008	7.020
Immig. noncoll.	0.199	0.176	0.007	3.184
Immig. coll.	0.212	0.235	0.020	1.192
South				
Native noncoll.	0.415	0.420	0.004	1.439
Native coll.	0.422	0.402	0.009	2.219
Immig. noncoll.	0.469	0.487	0.010	1.888
Immig. coll.	0.480	0.464	0.023	0.682
West				
Native noncoll.	0.343	0.350	0.004	2.123
Native coll.	0.350	0.313	0.008	4.487
Immig. noncoll.	0.332	0.337	0.009	0.568
Immig. coll.	0.308	0.300	0.021	0.361

Table 57: Probability of region choice, female born in '50s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncoll.	0.243	0.237	0.003	2.275
Native coll.	0.230	0.274	0.006	6.801
Immig. noncoll.	0.201	0.183	0.006	2.966
Immig. coll.	0.215	0.248	0.016	2.105
South				
Native noncoll.	0.421	0.427	0.003	1.902
Native coll.	0.426	0.424	0.007	0.265
Immig. noncoll.	0.479	0.486	0.008	0.784
Immig. coll.	0.480	0.499	0.018	1.042
West				
Native noncoll.	0.336	0.336	0.003	0.057
Native coll.	0.344	0.302	0.007	6.319
Immig. noncoll.	0.320	0.332	0.008	1.602
Immig. coll.	0.305	0.252	0.016	3.293

Table 58: Probability of region choice, female born in '60s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncoll.	0.246	0.253	0.004	2.005
Native coll.	0.234	0.269	0.007	5.381
Immig. noncoll.	0.205	0.199	0.006	0.884
Immig. coll.	0.218	0.248	0.014	2.222
South				
Native noncoll.	0.423	0.420	0.004	0.588
Native coll.	0.427	0.434	0.007	0.983
Immig. noncoll.	0.481	0.429	0.007	7.031
Immig. coll.	0.478	0.498	0.016	1.288
West				
Native noncoll.	0.331	0.327	0.004	1.239
Native coll.	0.340	0.297	0.007	6.292
Immig. noncoll.	0.314	0.371	0.007	7.932
Immig. coll.	0.304	0.254	0.014	3.685

Table 59: Probability of region choice, female born in '70s