Marrying your job: matching and mobility with geographic heterogeneity*

Davide Alonzo†

September 12, 2022

Please click here for the latest version.

Abstract

This paper examines the effects of geographic heterogeneity in occupational returns on marriage market outcomes and the impact of family formation on the geographic allocation of labor. We document that geographically mismatched workers – those living in a location that pays relatively lower wages to their occupation – are less likely to marry and more likely to divorce. We develop and estimate a model of migration and family formation. We assess individual and aggregate implications of joint marriage and location choices through counterfactual experiments. We find that, in aggregate, the marriage-market amenity enhances productivity by attracting workers to high-return locations.

JEL Classification: J12, J16, J21, J31, J61

Keywords: Migration, Marriage, Divorce, Geography

*I thank my Ph.D. committee members, Giovanni Gallipoli, Henry Siu, Michal Szkup, and David Green, for their support and guidance throughout the realization of this paper. I also thank Florian Hoffmann, Jesse Perla, Angela Redish, and all the seminar participants at the Vancouver School of Economics for many constructive suggestions. This research was enabled in part by support provided by the Digital Research Alliance of Canada.

†Université de Montréal. 3150, Rue Jean-Brillant, Montreal, H3T 1N8, Canada. E-mail: davidealonzo03@gmail.com

1 Introduction

The idea that marriage and geographic mobility interact is not new in the economic literature. On the one hand, following the work of Mincer (1978), economists have recognized that the presence of family ties imposes constraints on the geographic mobility of households with consequences for the labor market outcomes of its members (Guler, Guvenen and Violante, 2011; Gemici, 2016). In this sense, by preventing the relocation of married workers towards geographic areas in which they could be more productive, marriage constitutes a constraint that affects the efficiency of the geographic distribution of the labor force. On the other hand, marriage markets are local amenities that affect the evaluation that workers attach to different locations. As a consequence, the geographic heterogeneity in marriage market conditions influences the migration choices of households (Compton and Pollak, 2007; Edlund, 2005; Gautier, Svarer and Teulings, 2010).

This paper examines the joint effect of these forces on both aggregate and individual outcomes through an equilibrium framework that takes into account the endogeneity of marital and migration choices, the endogenous nature of marriage market conditions, and the heterogeneity in the value that different agents attach to the marriage-market amenity. The framework developed in this paper gives us some insights into three fundamental questions: (i) how does marriage affect the mobility choices of married and single households? (ii) What are the consequences of this interaction for family formation? (iii) What are the implications for the geographic distribution of the labor force?

With respect to the first question, the existing literature has focused for the most part on understanding the effects of family ties on migration and, as a consequence, on the labor market outcomes of married households. Most of these papers take marriage decisions as given and compare the labor market outcomes of married individuals to singles with similar observable characteristics. Yet, the marriage-market behavior of observationally equivalent individuals can be substantially different across local markets, making the comparison between married and single individuals not as straightforward. In other words, the endogenous selection into marriage is an important aspect to consider. The logic behind this reasoning is as follows. Introducing the terminology used in the rest of the paper, geographically mismatched workers – i.e., workers facing large geographic wage differentials and large potential labor market gains from mobility – might find it too costly to commit to a marriage due to potential future location conflicts with their spouses.[1] If this logic is correct, geographically

[1] To give a concrete example, a software engineer working in San Francisco, CA would be considered to be well matched as they are working in one of the best locations for their occupation. However, the same engineer working in El Paso, TX would be geographically mismatched as El Paso does not provide the same opportunities to software engineers as San Francisco.

mismatched individuals should display lower marriage hazard rates on average. Moreover, being local in nature, marriage markets display substantial heterogeneity, with the pool of potential partners including both mismatched and well-matched individuals (those facing low potential gains from mobility). Given the choice, a mismatched individual should prefer to marry a similarly mismatched partner, as they might be more inclined to migrate than a well-matched one. A specular logic holds for well-matched workers. Simply put, conditional on marriage, one should observe positive assortative mating between partners based on the size of the potential gains from migration (i.e., the value of the migration option). Finally, marriages involving mismatched workers should also be more fragile: upon realization of a location conflict (e.g., an appealing job offer from a different city for one of the spouses), the labor market mismatch might be enough to push some individuals to break the marriage. Thus, geographically mismatched individuals should display higher divorce rates.

The first section of this paper provides novel empirical evidence consistent with these mechanisms. This analysis takes advantage of the geographic heterogeneity in the wages paid to occupations. For each occupation, we rank cities based on city-occupation fixed effects obtained from a Mincer equation. These rankings deliver a simple measure of geographic mismatch: conditional on their occupation, a worker in a badly ranked city could earn more by moving to work in a different city. We estimate a set of probability models and show that geographically mismatched individuals are, in fact, less likely to marry and more likely to divorce than the well-matched. In particular, we find that geographically mismatched men (women) are as much as 23% (6%) less likely to enter marriage each year and, if married, 18% (8%) more likely to divorce than those who are not geographically mismatched. We also document a little-known dimension of marital sorting: looking at newlyweds, we show that homogamous marriages, e.g., well-matched workers marrying similarly well-matched workers, are relatively more frequent after controlling for marriage market conditions. In other words, assortative mating aligns with the magnitude of the potential gains from migration.

While intriguing, this evidence cannot be interpreted as causal due to the endogenous nature migration.[2] Moreover, as highlighted in the previous discussion, any attempt to address the aggregate effects of the interaction between marriage and migration should explicitly take into account the endogeneity of these choices. For these reasons, in section 3 we develop an equilibrium model of marriage formation and migration. The model economy consists of

[2]To understand this it is important to stress that our definition of geographic mismatch is solely based on wages. Nevertheless, because of the endogeneity of migration, being geographically mismatched does not necessarily imply a higher willingness to migrate. Geographically mismatched individuals might be well happy to be mismatched if their idiosyncratic evaluation of local amenities is high enough. In equilibrium, in fact, one can expect geographically mismatched individuals to have higher evaluations of local amenities. This introduces a bias in our empirical results as these individuals would also be more willing to enter marriage.

geographically segmented marriage and labor markets, each characterized by search frictions. Surplus within households is allocated through Nash bargaining, where the outside option is a costly divorce. Marriage market conditions are endogenously determined by the migration decisions of the agents. The novelty of the model lies in the determination of marriage market conditions through endogenous migration decisions, and in their mutual interaction. A key contribution of this analysis is that, to the best of our knowledge, it is the first to allow for the joint endogenous determination of migration and marriage behavior. The model is estimated through indirect inference targeting features of the empirical observations, as described in section 4.

We use the model to analyze counterfactual scenarios and gain insights into the three questions laid down earlier. The experiments in section 5 suggest that marriage substantially reduces geographic mobility. In a world without marriage, the probability that households migrate at least once over their lifecycle is about two to three times higher than in the baseline. Moreover, we find that, by entering an early marriage, men (women) give up as much as 10% (8%) of wage growth because of reduced mobility.

Furthermore, by comparing a counterfactual scenario in which we shut down the migration channel to the baseline, we find that mobility considerations reduce the average yearly marriage rate by about 3.7%. This average hides significant heterogeneity. Without migration, the yearly marriage probability of geographically mismatched individuals increases by 14.0%, while for the well-matched it increases by only 1.4%.

Finally, to gauge the aggregate implications of marriage for the geographic distribution of the labor force, we analyze a counterfactual scenario in which the option to marry is absent. The counterfactual shows that marriage acts as an agglomeration force: in the baseline specification, the population is more concentrated in a few big cities than in the counterfactual. As measured by the Herfindahl-Hirschman index the geographic concentration of workers increases by roughly 70% on average for each occupation. To quantify the effect of this reallocation of the labor force on productivity, we use average labor income as a proxy. In doing so, the model suggests that the reallocation of labor induced by marriage produces an increase in the average labor income of 2.7%. About 40% of this effect can be explained by the fact that marriage introduces an incentive for singles to relocate to more productive areas because of the more favorable marriage market conditions that arise endogenously with migration. The rest is explained by the observation that marriage, which is more likely among geographically well-matched individuals, prevents them from moving away from high-productivity areas.[3] In this respect, we also show through a simple

[3] In this respect and while we do not explicitly look at patterns along educational lines, our findings are more in line with those of Compton and Pollak (2007), who argue that power couples (both spouses have

tax exercise that incentivizing marriage can strengthen the agglomeration effect of marriage increasing the geographic concentration of workers and the average income.

Related literature. For some time now economists have inquired about aspects of the interaction between the mobility and marital choices of households. One branch of this literature takes marriage as given. Following the work of Mincer (1978), economists have recognized that the presence of family ties imposes constraints on the geographic mobility of households with consequences for the labor market outcomes of its members. Since then, several researchers have studied the labor supply and migration decisions of married households. Costa and Kahn (2000) focus on the location choices of college-educated couples, arguing that the joint location constraint causes them to relocate disproportionately more into large cities with thicker labor markets. In a mostly theoretical contribution, Guler, Guvenen and Violante (2011) examine the joint job search problem of couples and find that, by restricting geographic mobility, family ties worsen the labor market outcomes of married individuals relative to those of singles.[4] In an attempt to quantify these effects, Gemici (2016) estimates a structural model of migration with a dynamic framework of intrahousehold bargaining; her findings suggest that, compared to singles, married agents experience a lower wage growth over their working life. Braun, Nusbaum and Rupert (2019) use this mechanism to explain the observed fall in the migration rates of married households. The main difference between these papers and our work is that we explicitly consider selection into marriage and the heterogeneity of local market conditions while analyzing the costs of marriage in terms of labor market outcomes which allows us to explore additional sources of heterogeneity.

A smaller portion of this literature tries to link the migration patterns of singles to marriage. Compton and Pollak (2007), for instance, suggest that educated single individuals are attracted to large cities by the greater availability of potential mates. Similarly, Edlund (2005) and Gautier, Svarer, and Teulings (2010) find that, in the Swedish and Danish contexts, singles are more likely to move to cities because of better marriage opportunities. The kind of mechanisms suggested by these papers are allowed within our framework and, in fact, we find that the additional incentive to the migration of singles plays an important role in determining the geographic distribution of labor.

This work also relates to several strands of the literature on families and local labor markets. First, it features intrahousehold bargaining in a search and matching framework.

a college degree) are more likely to be found in larger cities because they tend to be formed there, than with those of Costa and Kahn (2000), who argue that this happens because power couples are more likely to relocate to bigger cities.

[4]The double search problem had also been used by Frank (1978) to explain the gender wage gap.

As noted by Lundberg and Pollak (1996), bargaining models provide an opportunity to integrate the study of the intrahousehold distribution of resources with a search and matching model of the marriage market. Within this literature, the bargaining protocol usually falls within two categories.[5] A popular option, and our approach in this paper, is to model the household allocation decision using Nash bargaining. The combination of search frictions and Nash bargaining is widely used in applied macroeconomics to make sense of structural changes such as the decline of marriage rates and the increased labor force participation of married women. Examples of such applications are Mazzocco, Ruiz and Yamaguchi (2013), Greenwood, Guler and Knowles (2000; 2003), Caucutt, Guner and Knowles (2002), Gemici (2016), Knowles (2013), Goussé, Jacquemet and Robin (2017). An alternative approach integrates the collective model, as in Chiappori (1988; 1992), with a bargaining protocol in which external shocks prompt a renegotiation of the sharing rule (i.e., the Pareto weights).[6] Examples of the application of this framework are Voena (2015), Devereux and Turner (2016), Low et al. (2018), Shephard (2019). Our choice of Nash bargaining over the latter protocol is driven by practical considerations. The model is computationally very and Nash bargaining, delivers a simple solution to the bargaining problem. Moreover, the alternative approach would require keeping track of an additional state variable (the Pareto weight) which would add complexity while having limited advantages.

Search and matching is just one way of modeling the marriage market. For example, Choo and Siow (2006; 2007) suggest equilibrium models of marriage formation that resemble a standard Walrasian market in which the price of marriage (i.e., a transfer from one spouse to the other) adjusts to equalize the demand for some partner characteristics to their supply. An extension of this model, allowing for imperfectly transferable utility, is due to Galichon, Kominers and Weber (2019). Applications and extensions of this class of models include Choo (2015) and Gayle and Shephard (2019). These models display several desirable properties, but they are not easily nested into a stationary equilibrium model of migration like ours.

An influential part of the literature on marriage markets investigates the causes and effects of "assortative mating" (Becker 1973; 1974*a*; 1974*b*), a term that describes the tendency of individuals differing in education, physical capital, height, race, and other traits, to marry partners with similar characteristics. The most commonly studied dimension of assortative mating is education. Examples of papers discussing the causes and the implications of

[5] Although these are the most popular, some alternative bargaining schemes are also used (see for instance Aiyagari, Greenwood and Guner, 2000, and Gallipoli and Turner, 2011). Moreover, the unitary model of the household is still actively used in some applications, for example in Greenwood, Seshadri and Yorukoglu (2005), Attanasio, Low and Sánchez-Marcos (2008), and Greenwood et al. (2016).

[6] See Chiappori and Meghir (2015) for a survey on the theoretical developments and empirical applications of this model.

educational sorting are Cancian and Reed (1998), Fernandez and Rogerson (2001), Schwartz (2010), Eika, Mogstad and Zafar (2019), Greenwood et al. (2014; 2016), Chiappori, Salanié and Weiss (2017). A few studies discuss assortative mating on other dimensions such as age (Choo and Siow, 2006, Diaz-Gimenez and Giolito, 2013) or other demographics (Siow, 2015). Our paper sets itself apart from this literature in that it highlights a non-trivial dimension of assortative mating, insofar agents who face different geographic wage distributions (i.e., have dissimilar incentives to migrate) have a lower marital surplus from marrying each other relative to agents facing similar wage distributions. Unlike most of the literature, in which the characteristic that drives marital sorting is either innate or determined early in life, in our model the sorting attribute is endogenous (due to migration) and dynamically changing.

This paper also speaks to the literature on workers' mobility in response to different economic conditions across geographic areas. A key contribution in this literature (Kennan and Walker, 2011) suggests, through the estimation of a tractable econometric model, that interstate migration is largely influenced by income prospects, but also that preferences play a non-negligible role. Our model relates to Kennan and Walker's as it features wage differentials and idiosyncratic preferences as key determinants of migration. However, our model differs in that the framework is characterized by search frictions, whereas they use a discrete choice framework, à la McFadden (1974), to describe the migration problem. More recent work by Diamond (2016) studies the interplay between local labor demand changes and the availability of local amenities in determining observed geographic sorting of skills. Using a structural spatial equilibrium model estimated on Census data, she finds that changes in local labor demand triggered geographic sorting, while the increasing amenities (due to the inflow of college graduates) worked as a reinforcing mechanism. Methodologically, the approach used in our paper is closest to Kaplan and Schulhofer-Wohl (2017), since they also use the geographic specificity of the returns to working in certain occupations as the relevant driver of migration. The main difference between these studies and our work is that we take a simplified view of the labor market, assuming away issues like unemployment and labor supply. On the one hand, this modeling choice substantially reduces the computational burden. On the other, labor market dynamics other than migration are neither necessary for the proposed mechanism to work nor are they the focus of this paper.[7] Moreover, we identify labor markets with metropolitan statistical areas, which results in a greater level

[7] In our model, local labor market conditions function as distribution factors, entering intrahousehold bargaining through the outside option. Thus, the actual amount of labor supplied by the household members is irrelevant as long as the value of the outside option is not affected by this choice. In reality, labor supply can indeed dynamically affect the value of the outside option, for instance through human capital accumulation. Nevertheless, we believe this is at most a second-order effect, especially in a world with relatively high labor force participation on both sides of the marriage market.

of geographic detail (unlike most of the studies in this literature, which focus on states or census divisions).

The starting point of many papers, including ours, is the presence of geographic wage differentials. Research suggests that these differences are only partly, if at all, due to selection on unobserved traits and mostly due to city characteristics, such as population size, skill distribution, and industry composition. Moretti (2004*a*; 2004*b*; 2004*c*; 2012) relates these differences to the concentration of college graduates in each MSA, arguing that a high concentration generates human capital spillovers that increase the productivity of all workers. Beaudry, Green and Sand (2012) attribute these gaps to differences in the industrial mix of cities. They argue that the presence of high-wage industries increases the wage rate paid to workers employed in other industries, improving their bargaining position thanks to a better outside option. Another branch of the literature documents the presence of city-size premia. For example, Baum-Snow and Pavan (2012) estimate a model of on-the-job search that incorporates endogenous migration between cities of different sizes and find that most of the size premium can be explained by differences in returns to experience and differences in wage intercepts while finding some evidence of negative sorting on unobserved ability. Similar results are found using Spanish data by De la Roca and Puga (2017). Using a large panel of worker-level data from Britain, D'Costa and Overman (2014) also find evidence of an urban wage premium that increases with city size, but find no significant differences in returns to experience across cities. Finally, Deming and Kahn (2018) use data on job vacancies to argue that differences in the wages paid in the same occupation are caused by differences in skill requirements. Determining the causes of wage differentials is interesting, but outside the scope of our work. The only requirement for our analysis to be valid is that these differentials exist and that they are not purely the result of geographic selection on unobserved idiosyncratic ability.

2 Empirical Evidence

In this section, we provide some empirical evidence consistent with the idea that geographic wage differentials affect marriage market outcomes. First, we construct a measure of geographic mismatch for each worker in our sample based on their occupation and city of residence. This measure reflects the value of the migration option of a worker conditional on their occupation or, in other words, their potential labor market gains from moving. Secondly, we provide evidence that this measure is informative, showing that geographically mismatched workers are more likely to move and that migration is directed towards better-paying cities. Finally, we show that highly mismatched workers are less likely to marry and

more likely to divorce. Moreover, conditional on marrying, they are more likely to marry similarly mismatched workers.

It is important to stress that the regression results shown below cannot be interpreted as causal. Their purpose is to unveil key correlations between geographic wage differentials and the way agents act in the marriage market. Some of these correlations will be used as the auxiliary model for the indirect inference estimation (Gourieroux, Monfort and Renault, 1993) of the structural model developed in the next section.

2.1 Data

To carry out our analysis we make use of data from the American Community Survey (ACS) (Ruggles et al., 2020). In particular, we pool together 10 years of ACS data from 2008 to 2017. Our choice of data is driven by two main needs: (i) as it will be clearer in what follows, to perform the analysis at the desired level of detail we need a rich dataset; (ii) we need information on migration and family formation and dissolution. To the best of our knowledge, this is the only readily available dataset of sufficient size that contains all the information needed for our analysis. Most of the other alternatives are either too small in the cross-section to allow an analysis at the city level (e.g., the Panel Study of Income Dynamics) or lack all the necessary information (e.g., the National Longitudinal Survey of Youth, which lacks information on migration). Similarly, administrative data often lack the necessary demographic information, such as education, that is indispensable when discussing marriage markets.

We restrict the sample to include only white, single (including divorced) or married households of workers aged between 25 and 55.[8] Moreover, we include only households living in urban areas. The first restriction is imposed to avoid confounding effects due to differences in the labor market conditions faced by different racial groups.[9] Moreover, marriage markets are very segmented along racial lines. Table 1 shows statistics on the composition of the 71,981 new marriages observed in the dataset. In almost three-quarters of the total, both

[8] An individual is considered to be married if their reported marital status is "married." In considering marriage as the relevant family bond, we follow the literature (Lundberg and Pollak, 2013 cite marriage as opposed to cohabitation as a commitment mechanism that supports higher investments in children). The ACS sample allows us to separately identify non-married members involved in a sentimental relationship. Cohabiting couples might behave differently in comparison to those who are legally married because, for instance, there is more uncertainty about the future of the relationship. In appendix A, we argue that this is indeed the case.

[9] As it will be clear from the next section, the empirical analysis is based on a series of city rankings that are constructed from estimated city-specific wage premia. Potential geographic differentials in the degree of race-based discrimination in the labor market require the aforementioned rankings to be race-specific. Unfortunately, the size of the sub-samples of racial minorities is too small to reliably estimate those wage premia and perform the analysis that follows.

		Women			
		White	Black	Other	Total
Men	White	51,502	476	3,585	55,563
		71%	0%	5%	76%
	Black	1,143	4,789	461	6,393
		2%	7%	1%	10%
	Other	2,489	164	7,372	10,025
		4%	0%	10%	14%
	Total	55,134	5,429	11,418	71,981
		77%	7%	16%	100%

Table 1: Marriages by race.

		Women		
		No College	Some College	Total
Men	No College	7,798	8,473	16,271
		15%	17%	32%
	Some College	4,288	30,943	35,231
		8%	60%	68%
	Total	12,086	39,416	51,502
		23%	77%	100%

Table 2: Marriages by education.

spouses are white, while in only about 7% they are both black and about 10% are among other races. This leaves us with a mere 12% of interracial marriages. The table also shows that among white male (female) newlyweds, only about 7% (7%) of them are in interracial relations. This figure reaches 25% (12%) for black and 26% (35%) for the residual racial group. By comparison, as shown in table 2, the marriage market is much less segmented on educational grounds.

The age restriction is imposed to shield the analysis from retirement-related effects which, at the same time, avoids the high frequency of occupation changes of young workers to affect the results (Papageorgiou, 2013; Gervais et al., 2016; Menzio, Telyukova and Visschers, 2016).

Finally, we restrict our sample to include only individuals supplying a positive amount of work and earning some labor income since for the others we cannot observe an occupation.[10].

The geographic unit used in the analysis is the metropolitan statistical area (MSA) from the U.S. Office of Management and Budget.[11] The occupational system we employ is based on an aggregation of the classification developed by Dorn (2009). This system has been designed to maintain a high level of detail while ensuring that all occupations are sufficiently

[10]Restricting the sample to full-time workers does not affect the results significantly

[11]Based on the 2010 standards.

represented in all cities.[12] Details about occupations and a complete crosswalk with Dorn's classification can be found in appendix J and robustness checks using the latter classification are performed in appendix C.

After dropping some of the smallest cities, the final sample consists of 895,346 single households and 1,448,744 married households, living in one of 209 MSAs (to which we refer interchangeably as "cities") and working in one of 95 occupations.

2.2 A Measure of Geographic Mismatch

To measure the geographic mismatch, we construct for each occupation a ranking of cities based on the city-occupation fixed effects obtained from an estimated Mincer equation. A worker that, conditional on their occupation, resides in a low-ranked city will have, other things equal, more profitable opportunities to migrate compared to a similar worker residing in a high-rank city.[13]

First, we estimate the following wage equation

$$\log w_{i,c,o,t} = \alpha_{c,o} + \boldsymbol{\beta X}_{i,c,o,t} + \epsilon_{i,c,o,t} \tag{1}$$

where $w_{i,t,c,o}$ is the real hourly wage of worker i living in city c and working in occupation o in year t. Real wages are computed following Moretti (2013).[14] The coefficients of interest, $\alpha_{c,o}$, are the city-occupation wage premia, and $\boldsymbol{X}_{i,c,o,t}$ is a set of controls that include dummies for education, a quadratic function of potential experience for each occupation, and dummies for gender, marital status and year. [15]

We use the estimated values of $\alpha_{c,o}$ to rank cities across occupations. The resulting rankings capture the quality of the city-occupation match.[16] One potential issue with the estimated fixed effects is selection bias.[17] Nevertheless, this is not a major obstacle for

[12] For this very reason, we exclude mining occupations that are extremely concentrated in particular regions.

[13] In this respect, our approach is akin to that of Kaplan and Schulhofer-Wohl (2017), who exploit changes in the dispersion of location-occupation premia to study migration flows.

[14] A local CPI index is calculated starting from the national CPI. We adjust the latter by allowing housing costs to vary by metropolitan area. We assign a fraction of the typical consumption basket to housing (this is stably around 41% during the sample years as reported by the Bureau of Labor Statistics) and adjust this component of the CPI using local housing prices obtained from the dataset as the average of the variable *rentgrs* for households dwelling in 2 or 3 bedroom houses.

[15] Education levels are high-school graduates and below, and some college and above. Potential experience is computed as (age - years of schooling - 6).

[16] See appendix B for more details about the distribution of the estimated fixed effects and the rankings.

[17] We cannot apply, in this context, the standard correction procedure developed by Dahl (2002). His technique does not allow to separately identify the intercept of the wage equation from the intercept of the control function, which makes cross-city comparisons unfeasible. In other words, it would allow us to control for the effects of selection when comparing two occupations within a city but not when comparing the same

Occupation	Low mismatch (1)	High mismatch (5)
Financial managers	new york-newark-jersey city bridgeport-stamford-norwalk san francisco-oakland-hayward	albuquerque el paso lafayette
Programmers	san jose-sunnyvale-santa clara san francisco-oakland-hayward seattle-tacoma-bellevue	jackson naples-immokalee-marco island clarksville
Lawyers and judges	washington-arlington-alexandria new york-newark-jersey city los angeles-long beach-anaheim	palm bay-melbourne-titusville tucson waco

Table 3: This tables displays some of the best and worst cities (in terms of average real wage rate) for a selection of occupations.

our empirical analysis. All the empirical results are, in fact, robust to selection as long as correcting for it would not completely explain the city premia, an eventuality that is not supported by the literature (Moretti, 2004*b*,*c*; D'Costa and Overman, 2014; Dahl, 2002; Dauth et al., 2019), or completely change the rankings.[18]

Given these rankings, we define a worker's geographic mismatch as follows: a worker that resides in a city that is among the 20% highest-ranked cities for their occupation is assigned a mismatch level of 1; a worker living in a city that is ranked among the next 20% of cities is assigned a mismatch of 2 and so on. In total, there are five levels of mismatch that capture the quality of the geographic match given one's occupation. Here, a level-5 worker displays the highest degree of geographic mismatch since they can earn a higher labor income simply by moving to almost any other city. Similarly, level-1 workers are geographically well matched. Table 3 shows some examples for a selection of occupations.

2.3 Migration, Marriage, Divorce and Geographic Mismatch

We now describe how our measure of geographic mismatch correlates with observed outcomes regarding migration, marriage, and divorce found in the data. Most of the analysis relies on estimating probability models with a similar structure. Let $Y_{i,c,o,t}$ be the outcome variable for household i living in city c, working in occupation o in year t. The general form of all the upcoming regressions is given by

occupation in two different cities.

[18]All the result will depend on the ordinality and not the cardinality of the fixed effects. Moreover, in the empirical application, the rankings will be divided into quintiles such that the results are robust to small variations in the rankings.

$$Pr\left(Yi, c, o, t = 1\right) = \varphi\left(\boldsymbol{\beta}_0 \boldsymbol{\gamma}_{i,c,o,t} + \boldsymbol{\beta}_1 \boldsymbol{X}_{i,c,o,t}\right) \tag{2}$$

where $\varphi(\cdot)$ is the logistic function, $\boldsymbol{\gamma}_{i,c,o,t}$ is a set of dummies corresponding to the 5 levels of geographic mismatch defined above or to all the possible combinations of the spouses mismatch in a married couple, depending on the application. Finally, $\boldsymbol{X}_{i,t,c,o}$ is a collection of controls that generally contains demographic variables and time dummies. It may also contain city characteristics and additional controls. These regressions are meant as evidence that the presence of potential gains from migration affects marriage and divorce behaviors. The upcoming results, though, cannot be interpreted as causal. The model developed in the section 3 will provide the means to rationalize these findings within the proposed mechanism.

2.3.1 Migration

We start by showing that the measure of geographic mismatch defined in the previous section is informative of the migration patterns observed in the data. The ACS surveys contain information regarding the city of residence in the year preceding the survey. We use it to construct a dummy variable that equals one if the current city of residence is different from the previous one. First, we consider single households and estimate the aforementioned logit model where the $Y_{i,t,c,o} = 1$ if the household has moved to a different city in the previous year to determine how the probability of migration changes with the worker's geographic mismatch. Clearly, the latter will generally change with the move; thus, in the regression, we use the mismatch level that characterized the individual before the move occurred. In the baseline specification, the vector of controls includes a dummy for the presence of children, a quadratic function of age, and a set of dummies for education and year. The analysis is carried out separately for men and women. The results are very similar for the two groups, thus we report here only the results for men and refer the reader to appendix B for the estimates for women and additional robustness.

The left panel of figure 1 shows the average probability of migration as a function of the five levels of mismatch. The graph shows that the probability of migrating is monotonically increasing in the geographic mismatch. For singles, the yearly migration rate is about 0.76%, but highly mismatched workers are more than twice more likely to migrate (1.32% vs. 0.57%). To get a sense of the direction of migration, the right panel in the same figure displays the negative average change of mismatch conditional on migrating.[19] The solid line shows the pattern we would observe if migration was random.[20] The fact that the estimated changes

[19] The negative sign is such that positive values correspond to an improvement.

[20] This assumes an equal probability of migrating to each city

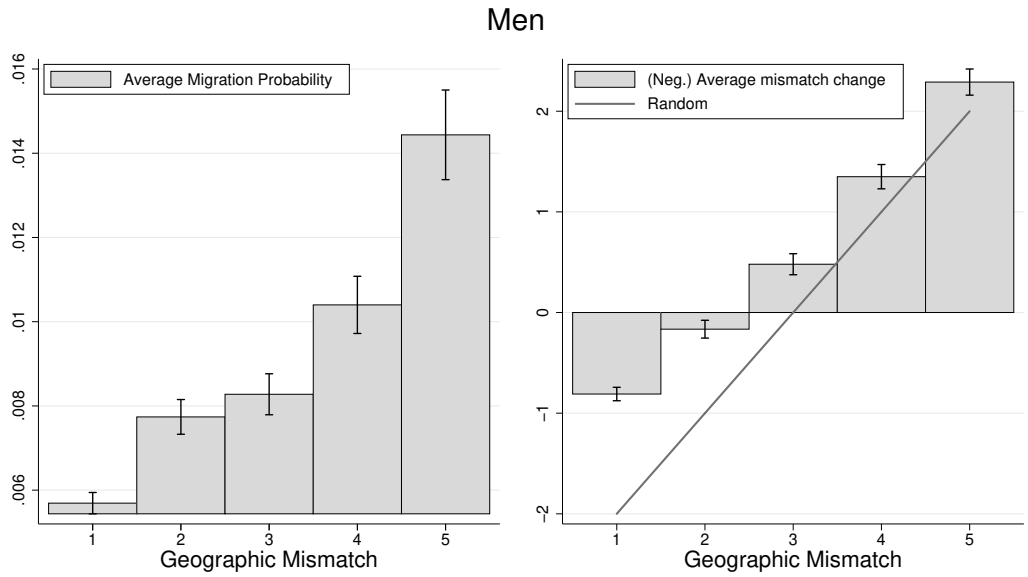

Figure 1: Migration rates and average (negative) change of geographic mismatch conditional on migration of single male households.

are always above the random case supports the idea that migration is directed toward cities that pay higher wage premia conditional on one's own occupation.

A similar pattern can be found for couples.[21] Figure 2 shows the (smoothed) average probability of migration for each combination of initial mismatch of husband and wife obtained from the estimation of a logit probability model of the type of equation 2, where the usual demographic controls are included for both husbands and wives. The figure shows that overall migration probabilities are increasing in the mismatch level of both husband and wife. Highly mismatched couples, with a probability of 0.4%, are twice as likely to move than a (1,1)-couple (because of smoothing, this difference appears to be smaller in the graph). Moreover, it is worth noticing that couples are overall about three times less mobile than singles, their yearly migration rate being 0.24%.

Occupational and Geographic Mobility. At this stage, it is important to discuss the issue of occupational mobility. Ideally, we would like to observe whether migration is associated with occupational changes or not as this would potentially entail a change in the level of geographic mismatch. Yet, our data do not contain any information on the occupational history of workers. We argue here that this is not detrimental to the validity of the results that follow.

[21] In the regression for couples, we drop households where husband and wife report having moved from two different cities.

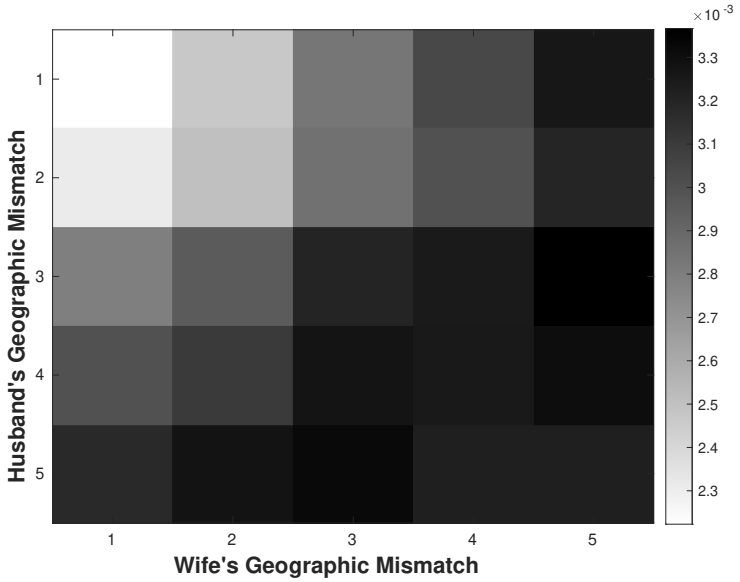

Figure 2: Married households migration rates as a function of the level of geographic mismatch of both spouses (smoothed).

Occupation switchers conditional on mobility

	Same house	Moved within county	Moved within state	Moved between states	Total
Same occupation	94.7%	91.1%	86.1%	79.7%	94.2%
Different occupation	5.3%	8.9%	13.9%	20.3%	5.8%
Total	100%	100%	100%	100%	100%

Table 4: Migration and job switching (2008-2017 CPS-ASEC).

First, we can observe from alternative data that over a one-year horizon (the frequency at which the upcoming analysis is performed) occupational switching is not a pervasive phenomenon. Table 4, constructed using CPS-ASEC data (Flood et al., 2020) from 2008 to 2017 applying the same sample restrictions as in the ACS sample, reports the fraction of workers who have changed occupation in the year preceding the survey conditional on migration. It shows that the vast majority of moves are not accompanied by occupation changes, with the overall fraction of occupational switches averaging 5.8%.

Secondly, the mechanism proposed in this paper can still be valid with respect to those who change occupations. If a worker is planning to change occupation and there are search frictions in the labor market, then the only relevant occupation for the proposed mechanism to work is the intended occupation and not the current one. In other words, a forward-looking individual who is planning to change occupation will consider the costs and benefits of geographic mobility associated with the intended occupation when making decisions regarding family formation. This implies that the relevant occupation is the current one and

not the one preceding the migration.

Finally, one can advance an equilibrium argument. Occupation mobility is costly and the cost of switching between two occupations depends on the similarities between them.[22] By a simple no-arbitrage argument, the returns between similar occupations should be very similar within a labor market and, thus, so should be their city rankings. If this is the case, the possibility of switching between occupations becomes irrelevant for the empirical analysis that follows, as most of the switches will occur between occupations with similar rankings.

2.3.2 Marriage

After having established the validity of our measure of geographic mismatch, we show evidence suggesting that the latter affects marriage market outcomes. First, we show that geographic mismatched workers are less likely to marry. Secondly, we provide evidence that, conditional on marrying, there is positive assortative mating on geographic mismatch (agents tend to marry similarly mismatched partners) and that this relationship is stronger the higher the level of geographic mismatch. As additional evidence of this assortative mating, we show that, after controlling for local marriage market conditions, the probability of marrying within an occupation is increasing in a worker's mismatch, meaning that geographically mismatched individuals are more likely to marry within their occupation.

In our sample, about 8.4% (9.4%) of single men (women) married within the year preceding the survey.[23] Using the information on new marriages, we estimate a linear probability model of the usual form for men and women. Here, the outcome variable $Y_{i,t,c,o}$ equals one if individual i married within the last year. The baseline specification includes controls for education, wage, city size, local sex ratios, the geographic dispersion of occupation-specific wage premia, year fixed effects, and a quartic in age. Tables and robustness checks are available in appendix B.

The estimation results are shown in figure 3. The left panel shows the results for men. The yearly probability of marriage is decreasing with geographic mismatch. A level-5 man is roughly 23% less likely to marry within a year than a level-1. For women, a smaller drop in this probability is visible only for the two highest mismatch levels, but only the 6% drop in the probability for those with the highest mismatch level is significant at 10%.

To get a sense of the magnitudes and of the cumulative effects of such differences, figure 4 shows a back-of-the-envelope calculation for the probability of marrying within a certain

[22] See Poletaev and Robinson (2008), Lazear (2009), Yamaguchi (2012) and Gathmann and Schönberg (2010).

[23] These statistics are calculated by dividing the number of newlyweds by the sum of the number of singles and that of newlyweds.

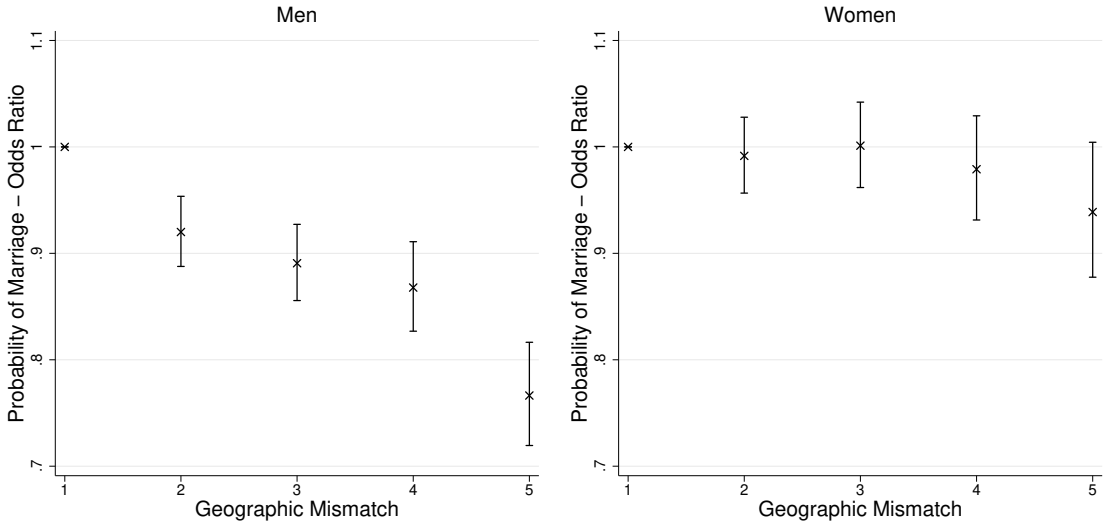

Figure 3: Yearly probability of marrying for different levels of geographic mismatch (odds ratio relative to level 1 with 95% confidence intervals). The average probability of marrying is 8.4% for men and 9.4% for women.

age obtained from the probabilities predicted by the estimated model for men.[24] According to this simple calculation, there is a sharp divergence between the marriage probabilities, and by age 32 the probability of having married at least once is more than 10 percentage points lower for level-5 men than it is for a level-1, a sizable difference. The corresponding figure for women can be found in appendix B.

To measure assortative mating, we borrow from Eika, Mogstad and Zafar (2019) and measure it by comparing the observed distribution of marriages over the space of possible mismatch levels to the one obtained randomly matching husbands and wives. Their measure of assortative mating adapted to the current problem, is

$$\tilde{s}_{h,w} = \frac{P(h,w)}{P(h)P(w)} \tag{3}$$

where h and w indicate the geographic mismatch of the husband and the wife respectively. This measure compares the observed distribution of new marriages across mismatch levels (numerator) to the one we would observe if matching was random (denominator). If $\tilde{s}_{h,w} > 1$ we have positive assortative mating while $\tilde{s}_{h,w} < 1$ is interpreted as negative assortativeness. The advantage of using this measure over more common measures, such as correlations, is

[24] For each observation i, we predict the average probability of marriage at different ages $\hat{p}_i^a$. The probability of marrying exactly at age $\bar{a}$ is given by $\tilde{p}_i^{\bar{a}} = \left[\prod_{a<\bar{a}} (1-\hat{p}_i^a)\right] \hat{p}_i^{\bar{a}}$. Thus the estimated probability of marrying before age a is computed as $p_i^a = \sum_{\bar{a}\leq a} \tilde{p}_i^{\bar{a}}$. The values reported in figure 4 are simply obtained as the averages of p_i^a for the specified level of geographic mismatch.

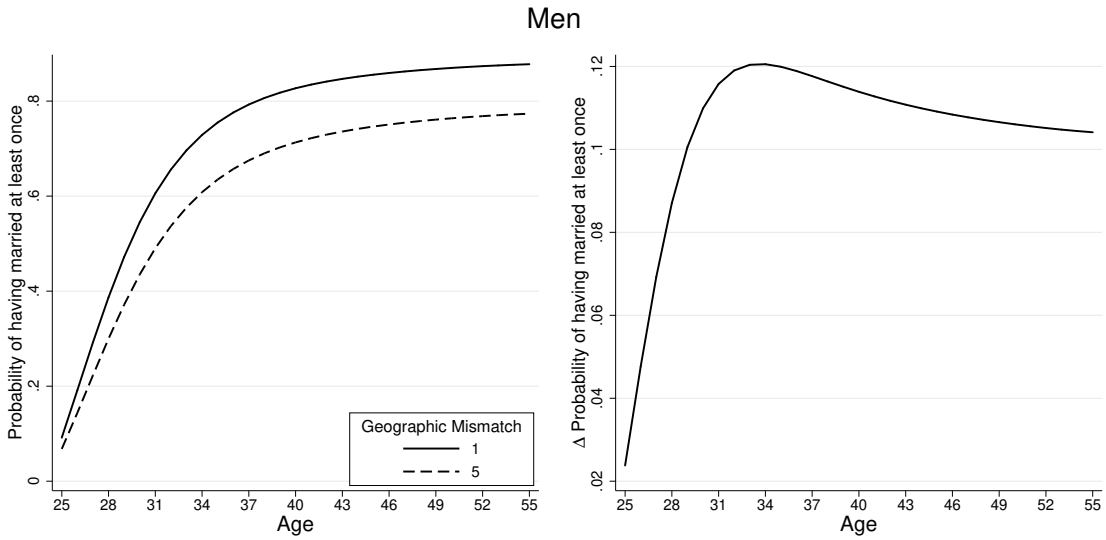

Figure 4: Estimated probability of having married at least once as a function of age for men with the lowest and highest levels of geographic mismatch. The left panel shows the estimated probabilities while the right panel the difference between the two.

that it allows us to capture the "intensity" of assortative mating for each specific pair of mismatch levels.[25]

One of the shortcomings of the latter measure is that it does not take into account the local nature of marriage markets. To take the latter into account, we enhance our measure of marital sorting allowing the counterfactual random matching to occur only within cities. This is done to control for the composition of the local marriage markets because bigger and more productive cities tend to have a bigger fraction of well-matched workers compared to smaller cities. Our preferred measure of assortative mating is then

$$s_{h,w} = \frac{P(h,w)}{\sum_c P(h|c)P(w|c)P(c)} \tag{4}$$

where the numerator is the probability of observing a couple with mismatch level (h, w) and the denominator is the probability of observing the same type of couple in the counterfactual random distribution. The latter is equal to the sum of the probabilities of observing such a couple in each city c under random matching, $P(h|c)P(w|c)$, weighted by the fraction of couples living in the city $P(c)$.

We compute these measures for new couples, to avoid the confounding effects of later migration. Appendix B reports some robustness checks.[26] Figure 5 shows the results of these

[25]For a discussion on different measures of assortative mating, see Chiappori, Costa-Dias and Meghir (2020).

[26]We compute the same measure including all couples and compare it to the one in the main text.

computations, where the top panels refer to the "naive" measure, $\tilde{s}_{h,w}$, and the bottom panels to the preferred measure, $s_{h,w}$, which controls for local marriage market conditions. The left panels plot the computed values for each combination of mismatch levels. Clearly, the graph shows that there is positive assortative mating within mismatch levels and negative assortative mating between them, this is true for both measures. The right panels of the figure focus on the diagonal elements, i.e., it shows the two measures for $h = w = x$ and includes bootstrapped standard errors. Interestingly and in accordance with our conjecture, the degree of assortative mating is increasing in the geographic mismatch. Moreover, it is worth noticing that the differences between the top and bottom panels highlight the importance of considering the local nature of marriage markets. Despite providing comparable qualitative results, without controlling for local market conditions we would overestimate the degree of marital sorting, both positive and negative.

A consistent part of the homogamous couples in the sample is composed of couples in which husbands and wives work in the same occupation. Among new couples, this fraction is 8.21%. Although this is not in contrast with the theory, it raises the question of whether the observed mating patterns are driven by reasons other than the claimed interaction of marriage decisions with the potential pecuniary gains from migration. In the latter case, under the assumption that these other reasons are independent of geography, we should observe that the probability of marrying someone in the same occupation is independent of the degree of geographic mismatch after controlling for personal characteristics and conditions in the local marriage market. To test this hypothesis, we estimate the usual logit model in equation (2), where the outcome variable is being in an occupationally homogamous marriage, on the subset of newlyweds (separately for men and women). The vector of covariates $\boldsymbol{X}_{i,c,o,t}$ includes, in the baseline specification, a quartic in age, year dummies, and controls for the fraction of workers of the opposite sex working in the same occupation as individual i, for the education level of both spouses (with interaction) and the wage earned by both spouses (with interaction). Figure 6 plots the estimated odds ratios for the five levels of geographic mismatch. There is a clear upwards sloping profile for men. Highly geographically mismatched men are, conditionally on marrying, more likely to marry a woman in the same occupation than well-matched individuals by about 50%. Not surprisingly, the estimates for women closely mirror that of men.

Moreover, the bigger sample size allows us to extend the measure in (4) to control for other demographic characteristics like age and education.

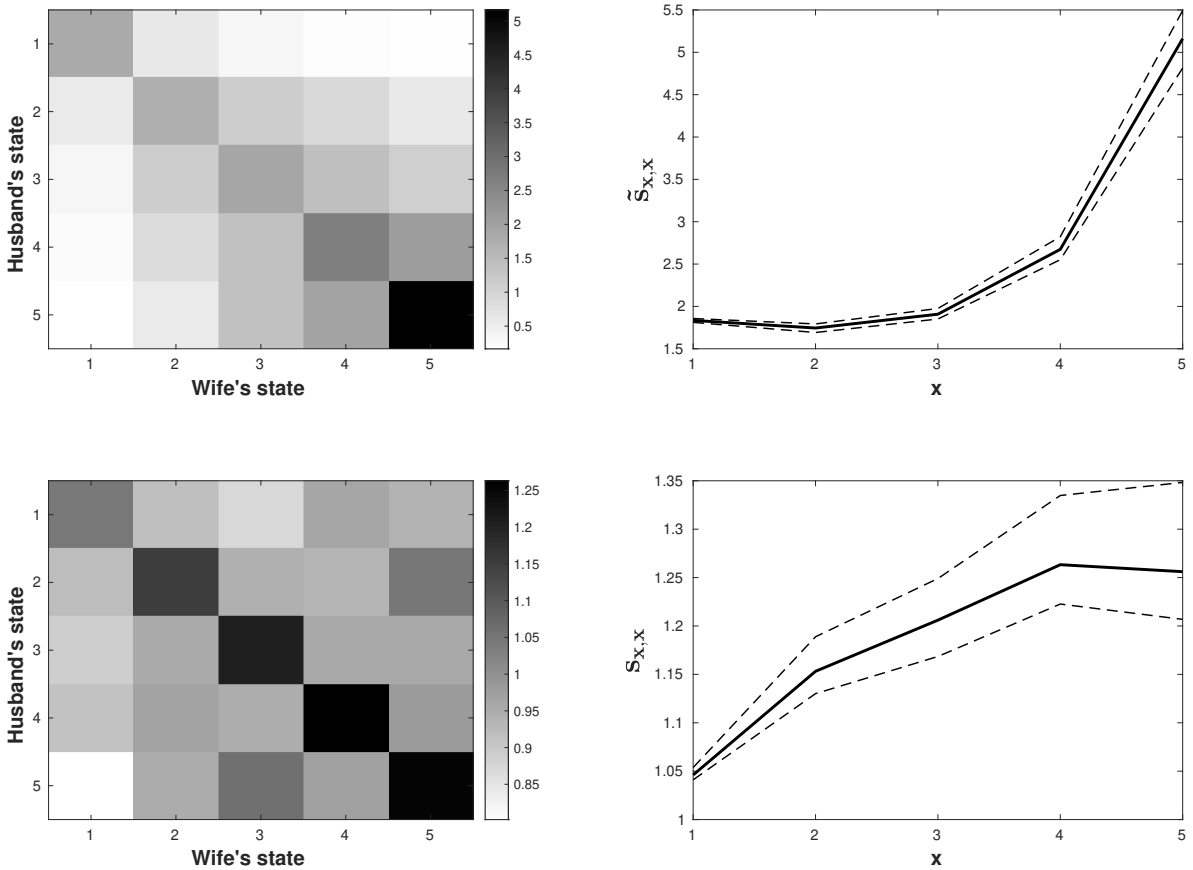

Figure 5: Measures of marital sorting as a function of the spouses' geographic mismatch. The left panels show all the values, while the right panels focus on the diagonal values and include 95% confidence intervals. The bottom two panels show the results for the preferred measure, $s_{h,w}$, which controls for local marriage market conditions.

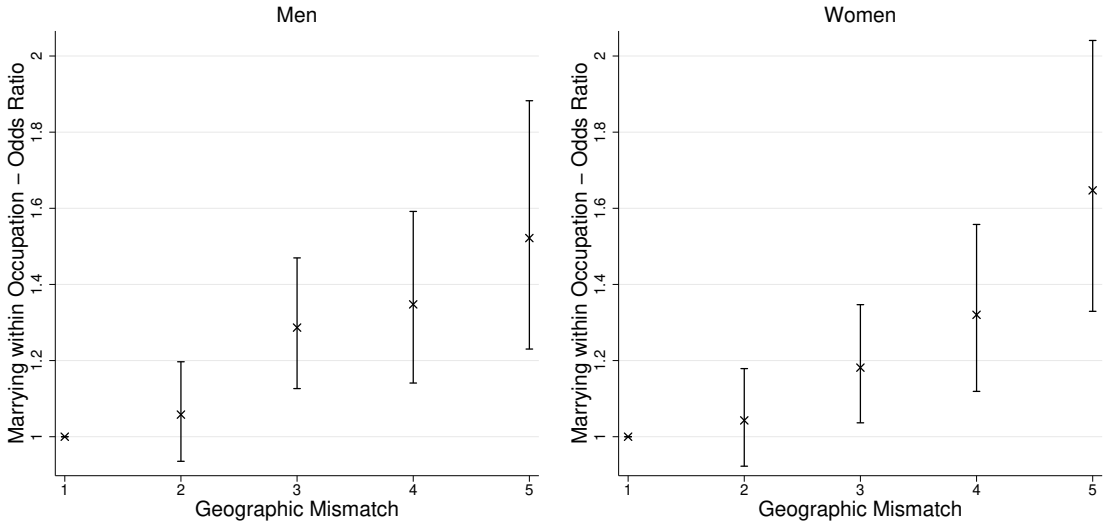

Figure 6: Probability of having married within an occupation as a function of geographic mismatch (odds ratios). The left panel reports the results for men and the right panel for women. The fraction of newlywed couples who work in the same occupation is 8.5%.

2.3.3 Divorce

The last piece of evidence we provide regards divorce. Theoretically, by choosing to divorce, individuals do not face the double search problem anymore. Given the higher value of the migration option, the outside option of marriage, divorce, is on average more attractive to geographically mismatched spouses. This implies that the surplus from marriage is lower and, thus, marriages are less stable. In other words, we expect highly mismatched individuals to be more likely to divorce than the well matched. To investigate this hypothesis, we estimate equation (2) on the subset of married individuals and divorcees. Here, $Y_{i,c,o,t} = 1$ if individual i has divorced within the last year, and $\boldsymbol{X}_{i,t,c,o}$ controls for education, the presence of children under 5, wage, age (quartic), the geographic dispersion of occupation-specific wage premia and year fixed effects. As for the marriage regression, we report in figure 7 the estimated odds ratios for the different levels of mismatch. Tables and robustness are in appendix B. For both men and women, we observe a higher probability of divorce for geographically mismatched spouses. The effect is stronger for men as the probability of divorcing within a year is more than 10% higher for level-5 men than it is for level-1 men. For women, the difference is above 5%. Overall, the average probability of divorce is 2.7% for men and 3% for women.

Additional evidence in support of the theoretical link between divorce and migration can be found by analyzing the migration patterns of divorcees, and comparing them to that of

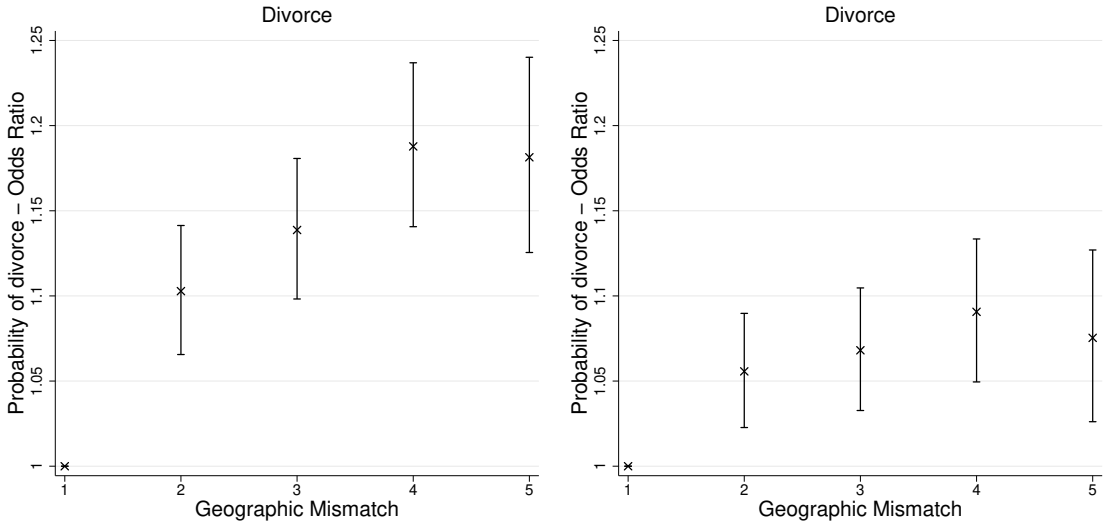

Figure 7: Average probability of divorcing as a function of geographic mismatch (odds ratios). The left panel reports the results for men and the right panel for women. The average probability of divorce is 2.7% for men and 3% for women.

	New Divorcees vs. Singles	All Divorcees vs. Singles	New Divorcees vs. Married	All Divorcees vs. Married
Men	1.95	1.37	3.46	2.38
Women	2.19	1.41	4.18	3.28

Table 5: Regression-based odds-ratios for migration comparing different groups. The estimated logit models contain controls for age, wages, geographic dispersion of occupation-specific wages, presence of children, and education. All the estimated odds-ratios are significative at 0.1%.

singles and married individuals. To perform this comparison, we estimate a series of logit models for the probability of migration of the following form on different subsets of the population:

$$Pr\left(Yi, c, o, t = 1\right) = \varphi\left(\beta_0 + \beta_1 Divorcee_{i,c,o,t} + \boldsymbol{\beta}_2 \boldsymbol{X}_{i,c,o,t}\right) \tag{5}$$

where $Divorcee_{i,c,o,t}$ is a dummy indicating whether individual i is a divorcee and $\boldsymbol{X}_{i,c,o,t}$ is a set of controls for age, wage income, the geographic dispersion of occupation-specific wages, the presence of children, and education.

Table 5 shows the estimates for the coefficient on the dummy for divorce in odds-ratio form for four different subsets: (i) singles (never married) and recent divorcees; (ii) singles and all divorcees; (iii) married and recent divorcees; (iv) married and all divorcees. The logit models are estimated separately for men and women. The results clearly show that,

after controlling for observables, the divorcees are more likely to migrate than both married individuals, and this difference is the largest when we consider only recent (less than a year) divorcees. This is in line with the idea that marriage is restricting the migration possibilities of spouses, and that these restrictions are lifted with divorce. The results also suggest that divorcees are more likely to migrate than singles, which is also in line with the hypothesis that marriage is restricting the mobility of spouses. Theoretically, single individuals, being able to migrate freely, are more likely to have already moved to their preferred city. The same is not true for recent divorcees. For this reason, when compared to singles, and after controlling for observables, they are found to be more likely to migrate than singles as well.

As for singles, we also compute the average change in geographic mismatch for recent divorcees. As shown graphically in figure 23 in appendix B, these patterns are comparable to those of singles.

3 Model

In this section, we develop a search and matching model of marriage and divorce where endogenous migration flows affect marriage outcomes through their effects on the composition of marriage markets and, at the same time, where migration flows are affected by marriage as forward-looking individuals take marriage market conditions in consideration when deciding upon migration opportunities. We consider a perpetual youth model (Blanchard, 1985) in which time is discrete. Workers discount the future at rate $\beta = \omega\tilde{\beta}$, where $\tilde{\beta}$ is the intertemporal discount factor and ω is the survival probability (common to all agents). They live in one of C cities and work in one of J occupations. Each agent is characterized by their occupation j_g, their city of residence c, a preference parameter for the current residence ξ_g, and their marital status (single or married). The index $g = m, f$ indicates the gender. Married households are also characterized by an additional state capturing the marriage match quality ("bliss" parameter) ζ. In the model, there is no unemployment, and the wages received by workers are exogenously determined solely by their city of residence and occupation. We restrict our attention to stationary equilibria. Moreover, we assume that in married households husband and wife die together and that, upon death, each agent is replaced by a newborn single worker of the same sex and type drawn randomly from some distribution. Finally, we allow workers to change occupation randomly according to an exogenous transition matrix.[27]

[27] Allowing for endogenous occupational switching would require adding to the already onerous model a full theory of occupational choice. Although it might be interesting to study the interplay between marriage, migration, and occupation choices, that is not the goal of this paper, and adding this extra layer of complexity would be computationally infeasible and would add little to the current analysis.

In what follows, we adopt a convenient within-period timing structure. At the beginning of each period, the marriage market operates. Singles are randomly matched to individuals of the opposite sex living in the same city (i.e., marriage markets are local). Upon matching, singles observe the characteristics of their match, draw a value for the bliss parameter of the match, and decide whether to get married or not. In this phase, couples simply update match quality ζ which follows an exogenous stochastic process.

Following the marriage market, the labor market operates. In each period, households are hit with some probability by a shock that allows the household to relocate to another specific city. We call this a "mobility shock". Given this shock, single households decide whether to move or not, while married households have to choose among a set of options: (i) stay married and move to the new city; (ii) stay married and do not move; (iii) divorce and one of the spouses stays and the other moves; (iv) divorce and both move; (v) divorce and both stay.[28] Divorcees re-enter the marriage market as singles in the following period.

After migration decisions are taken, workers receive their wages. We assume there is no saving technology, and households consume all their income and enjoy city amenities. Finally, right before the end of each period and after consumption, all individuals receive a shock to their preference for the current city and may stochastically change occupation. The timing structure is summarized in Figure 8.

We assume that the allocation of consumption within a married household is determined by Nash bargaining. The main reason for this modeling choice is that, together with the assumption of linear utility, it implies perfectly transferable utility and substantially simplifies the solution of the intrahousehold allocation problem, reducing the already heavy computational burden.

3.1 Marriage Market

In this section, we present the Bellman equations associated with the marriage market problem of single and married women. The value functions for men are symmetrical. At the beginning of each period and after the preference shocks are realized, singles and couples enter the marriage market.

Couples. In this phase, couples simply update their match quality parameter. We assume that divorce decisions are postponed to the labor market phase, i.e., divorces can occur only after the mobility shock is realized. This assumption rules out inefficient divorces as

[28] In the model the latter two options can be optimal only after a negative shock to the marriage quality parameter ζ. Postponing the divorce decision until after mobility opportunities materialize prevents inefficient divorces within the period.

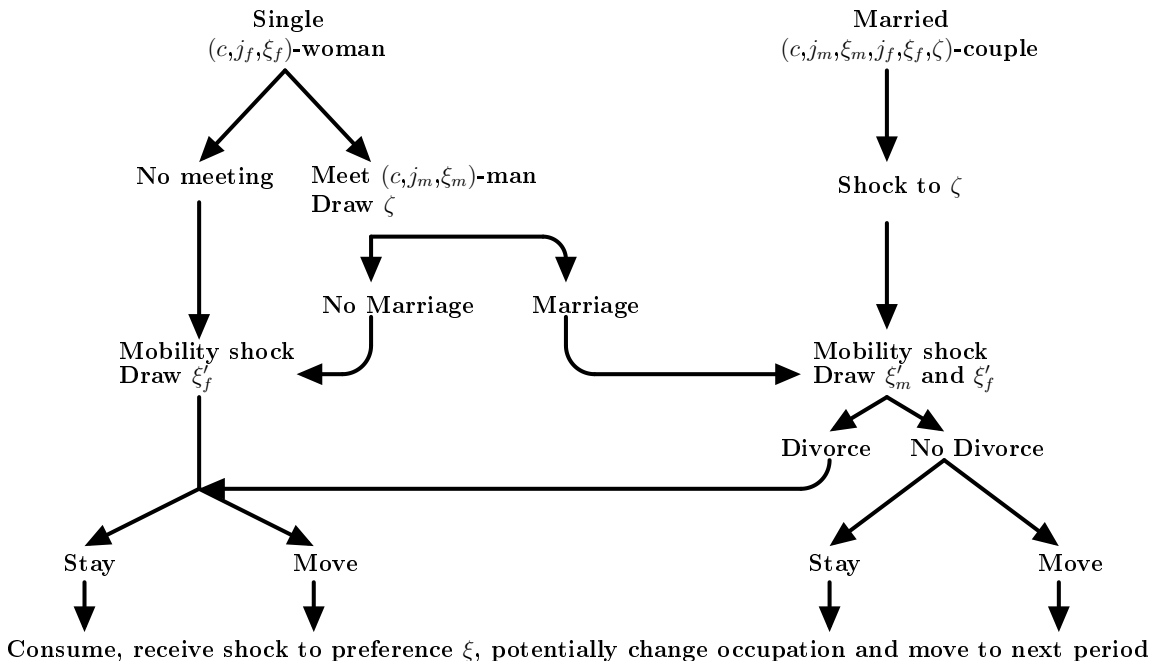

Figure 8: Within-period timing structure.

it prevents the break-up of marriages that might have survived after migration. We also assume that ζ evolves according to an autoregressive process given by

$$\zeta' = (1-\rho_\zeta)\,\bar{\zeta} + \rho_\zeta \zeta + \epsilon_\zeta \qquad \epsilon_\zeta \sim N(0, \sigma_\zeta) \tag{6}$$

Let $f_\zeta(\zeta'|\zeta)$ be the conditional distribution of ζ' given ζ. Moreover, to save on notation, define $x_m = \{j_m, \xi_m\}$ and $x_f = \{j_f, \xi_f\}$. The value for a married woman at the beginning of each period is

$$V^m_{f,1}(c, x_m, x_f, \zeta) = \int_{\zeta'} V^m_{f,2}(c, x_m, x_f, \zeta')\, f_\zeta(\zeta'|\zeta) \tag{7}$$

where $V^m_{f,2}$ is the value of being a married woman right before the labor market phase.

Singles. In a search and matching framework, a single woman of type x_f has a positive probability, g_{f,c,x_f,x_m}, of being matched to a type x_m man living in the same city. The probability of matching depends on the availability of singles in the city of residence and is determined as follows

$$g_{f,c,x_f,x_m} = \frac{M_{x_m,x_f}}{\mu_{c,f,x_f}} \tag{8}$$

where μ_{c,f,x_f} is the mass of single women of type x_f living in city c and M_{x_m,x_f} is the total number of matches occurring between type-x_m men and type-x_f women in city c. In

particular, we assume that

$$M_{x_m,x_f} = \underbrace{\lambda\left(1+\gamma\mathbb{1}\{j_m = j_f\}\right)}_{A}\underbrace{\mu_{c,m}^{\alpha}\mu_{c,f}^{1-\alpha}}_{B}\underbrace{\frac{\mu_{c,m,x_m}}{\mu_{c,m}}\frac{\mu_{c,f,x_f}}{\mu_{c,f}}}_{C} \tag{9}$$

The matching function displays constant returns to scale and is composed of three terms.[29] The first (A) is the overall matching efficiency. If the coefficient γ is bigger than zero, matches within occupations are more likely. The presence of this bias in the matching function is needed for the model to quantitatively match the fraction of marriages occurring between workers in the same occupation.[30] The second term (B) captures how the number of matches is affected by the overall balance on the two sides of the market. Taken together and ignoring the γ term, A and B represent the standard Cobb-Douglas matching function widely used in the search literature. The last term (C) implies that the total number of matches of any type is proportional to the shares of single men and women of the same types that populate the local economy.[31]

Upon matching, a match quality parameter is drawn from the stationary distribution of ζ, $f_\zeta(\zeta)$. Let μ_{c,m,x_m} be the mass of single men of type x_m in city c (thus $\mu_{c,m} = \int_{x_m}\mu_{c,m,x_m}dx_m$), the value of a single woman at the beginning of each period is

$$\begin{aligned} V_{f,1}^{s}(c,x_f) = \left(1-\int_{x_m} g_{f,c,x_f,x_m}dx_m\right)V_{f,2}^{s}(c,x_f) + \int_{x_m} g_{f,c,x_f,x_m}\int_{\zeta}\Big[\left(1-m(c,x_m,x_f,\zeta)\right)V_{f,2}^{s}(c,x_f) \\ + m(c,x_m,x_f,\zeta)\,V_{f,2}^{m}(c,x_m,x_f,\zeta)\Big]\,df_\zeta(\zeta)\,dx_m \end{aligned} \tag{10}$$

where $m(c,x_m,x_f,\zeta)$ is the policy function for marriage and it equals one if marriage is profitable for both the man and the woman, i.e., if

$$\begin{aligned} V_{f,2}^{s}(c,x_f) \leq & V_{f,2}^{m}(c,x_m,x_f,\zeta) \\ V_{m,2}^{s}(c,x_f) \leq & V_{m,2}^{m}(c,x_m,x_f,\zeta) \end{aligned} \tag{11}$$

This implies that single women will follow a threshold rule and marry only if the drawn marriage quality is above some state-dependent level.

[29] Constant returns to scale in marriage markets are empirically supported by Botticini and Siow (2007).

[30] Even without this bias, the model does generate a relatively large fraction of occupationally homogamous marriages. Nevertheless, without a positive γ the model cannot match the empirical figure.

[31] It is worth noticing that, with $\gamma = 0$, the expression for the total number of matches occurring in a city takes the usual Cobb-Douglas form, namely

$$\sum_{x_m}\sum_{x_f} M_{x_m,x_f} = \lambda\mu_{c,m}^{\alpha}\mu_{c,f}^{1-\alpha}$$

3.2 Migration and Divorce

Here, we present the value functions associated with the migration problem. In this model, the labor market is very stylized. There is no unemployment and wages are exogenously fixed. This labor market reduces, then, to a location choice. We assume search frictions: in each period, households may receive a mobility shock that allows them to evaluate migration towards one randomly chosen city.[32] Upon receiving the shock, each household member draws a preference parameter for the candidate city from a continuous distribution, $f_\xi(\xi)$, and migration decisions are made. Importantly, this structure implicitly assumes that workers have no memory of past locations. At the beginning of each period, only the preference for the current city enters the state space. This preference is forgotten upon migration. This modeling choice is mainly driven by practical considerations as including memory, even limited, would severely increase the size of the state space and thus the computational burden.

Singles. Upon receiving the mobility shock and after drawing the preference parameter for the candidate city, the single household will migrate if the value of moving is greater than the value of staying. The moving cost is assumed to be a fixed fraction, κ, of income and it captures both pecuniary costs (e.g., shipping costs), the cost of temporary unemployment (since unemployment is not explicitly modeled), and psychological costs.[33] The corresponding Bellman equation is

$$\begin{aligned} V^s_{f,2}(c, x_f) = (1-\chi)\left[w_{f,c,j} + \xi_f + \beta E_{x'_f}\left[V^s_{f,1}\left(c, x'_f\right)\right]\right] \\ + \chi \sum_{c'|c'\neq c} \theta(c'|c) \int_{\tilde{\xi}_f} \max\Big\{ w_{f,c,j} + \xi_f + \beta E_{x'_f}\left[V^s_{f,1}\left(c, x'_f\right)\right], \\ (1-\kappa) w_{f,c',j} + \tilde{\xi}_f + \beta E_{x'_f}\left[V^s_{f,1}\left(c', x'_f\right)\right]\Big\} df_\xi\left(\xi'_f\right) \end{aligned} \quad (12)$$

[32] An alternative option is to model the migration choice to use a multinomial logit structure á la McFadden (1974) (see Kennan and Walker, 2011). The main advantage of such a structure is that, thanks to the properties of the extreme value distribution, migration choices would be fully characterized by a set of probabilities that can be directly computed from data. The problem with applying such a structure here is that, for spouses to bargain directly on each mobility option, it would be necessary to keep track of all the draws of the preference parameter of both spouses for each city and compute the bargaining solution for each combination of cities.

[33] This cost is independent of distance since we do not directly map model cities to real cities. We do not believe this to be a major issue since, according to CPS data, about 86% of moves occurred within state boundaries,

where $w_{f,c,j}$ is the wage paid to women working in occupation j in city c and $\tilde{\xi}_f$ is the preference for the candidate city. Moreover, $E_{x'_f}$ is the expectation operator, where the expectation is taken with respect to the preference parameter ξ and the conditional probability of occupational switching (i.e. it accounts for the preference shock and the possible occupation change occurring between two periods). The preference parameter ξ is assumed to evolve according to an AR(1) process:[34]

$$\xi' = \rho_\xi \xi + \epsilon_\xi \qquad \epsilon_\xi \sim N\left(0, \sigma_\xi\right). \tag{13}$$

The parameter χ captures the probability of receiving any mobility shock and $\theta\left(c'|c\right)$ is the probability of the shock coming from city c' given that our agent is currently residing in c (clearly $\sum_{c'|c'\neq c} \theta\left(c'|c\right) = 1$). In our application, we assume that the latter conditional probability is given by

$$\theta\left(c'|c\right) \equiv \theta = \frac{1}{C-1}. \tag{14}$$

Occupational switching follows a time-independent Markov chain with typical elements given by $\pi\left(j'|j\right)$.

The policy function associated with the migration choice in equation (12) is denoted with $t_f\left(c, x_f, c', \tilde{\xi}_f\right)$. The latter equals one if a single woman of type x_f finds it optimal to move from c to c' given that her preference for the latter city is $\tilde{\xi}_f$.

Couples. Couples allocate resources through Nash bargaining (with η capturing the bargaining power of the wife), where the outside option is divorce. If no mobility shock is realized, the surplus from marriage is given by

$$\begin{aligned} S_{nof}\left(c, x_m, x_f, \zeta\right) &= x\left(w_{m,c,j}, w_{f,c,j}\right) - \left(1-\delta\right)\left(w_{m,c,j} + w_{f,c,j}\right) + \zeta \\ &+ \beta E_{\boldsymbol{x}'}\left[\left(V^m_{m,1}\left(c, x'_m, x'_f, \zeta\right) - V^s_{m,1}\left(c, x'_m\right)\right) + \left(V^m_{f,1}\left(c, x'_m, x'_f, \zeta\right) - V^s_{f,1}\left(x'_f\right)\right)\right] \end{aligned} \tag{15}$$

where $x\left(a,b\right) = \left(a^\rho + b^\rho\right)^{\frac{1}{\rho}}$ captures economies of scale in consumption since $x\left(a,b\right) \geq \left(a+b\right)$ if $\rho \leq 1$. Moreover, δ is the proportional cost of divorce, paid separately by both divorcees, and $E_{\boldsymbol{x}}$ is the expectation operator with respect to the preference parameters and occupations of both husband and wife. Marriage survives if $S_{nof}\left(c, x_m, x_f, \zeta\right) \geq 0$. Thus, conditional on not receiving a mobility shock, married agents get their outside options plus the maximum

[34] Notice that the process has mean zero. This is just a normalization since the mean of ξ is irrelevant to the household's choices.

between their share of marital surplus (no divorce) and zero (divorce):

$$V^m_{f,nof}\left(c,x_m,x_f,\zeta\right)=\left(1-\delta\right)w_{f,c,j}+\xi_f+\beta E_{x'}\left[V^s_{f,1,t+1}\left(c,x'_f\right)\right]+\max\left\{0,\eta S_{nof}\left(c,x_m,x_f,\zeta\right)\right\}. \tag{16}$$

The policy function corresponding to the divorce choice described in the max operator is denoted by $d_{nof}\left(c,x_m,x_f,\zeta\right)$, and it equals one if divorce is optimal and zero otherwise.

If a mobility shock is received, six outcomes are possible: (i) the couple moves; (ii) the couple does not move; (iii) the couple divorces and only the husband moves; (iv) the couple divorces and only the wife moves; (v) divorce and both move; (vi) divorce and both stay. For options (i) and (ii), Nash bargaining is carried out independently with the outside options for husband and wife being divorced followed by the optimal individual choice about migration. This setup allows spouses who would privately gain from moving (staying) to make intratemporal transfer to a disagreeing partner in order to incentivize them to move (stay). The aforementioned transfer compensates the partner for all the expected lifecycle losses. Marriage survives, and the choice of the household is either (i) or (ii) as long as the marital surplus is positive.

Conditional on the mobility shock, the outside option for a woman is given by

$$\begin{aligned}&V_{f,out}\left(c,x_f,c',\tilde{\xi}_f\right)=\\&\max\left\{\left(1-\delta\right)w_{f,c,j}+\xi_f+\beta E_{x'}\left[V^s_{f,1}\left(c,x'_f\right)\right],\left(1-\kappa-\delta\right)w^F_{c',j}+\tilde{\xi}_f+\beta E_{x'}\left[V^s_{f,1}\left(c',x'_f\right)\right]\right\}\end{aligned} \tag{17}$$

where the first term in the max operator is the value of staying in city c and the second is the value of moving to city c'. The policy function associated with this choice is $t_{out,f}\left(c,x_f,c',\tilde{\xi}_f\right)$. Notice that this function is, in general, different from that of singles because of the presence of the proportional divorce cost. The marital surplus conditional on moving is

$$\begin{aligned}&S_{move}\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right)=x(\tilde{w}_{m,c',j},\tilde{w}_{f,c',j})+\tilde{\xi}_m+\tilde{\xi}_f+\zeta\\&+\beta E_{\boldsymbol{x'}}\left[V^m_{m,1}\left(c',x'_m,x'_f,\zeta\right)+V^m_{f,1}\left(c',x'_m,x'_f,\zeta\right)\right]-\left[V_{m,out}\left(c,x_m,c',\tilde{\xi}_m\right)+V_{f,out}\left(c,x_f,c',\tilde{\xi}_f\right)\right]\end{aligned} \tag{18}$$

where $\tilde{w}_{g,c,j}=(1-\kappa)w_{g,c,j}$ is income after having paid the moving cost. The surplus

associated with the choice of staying is

$$S_{stay}\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right)=x(w_{m,c,j},w_{f,c,j})+\xi_m+\xi_f+\zeta$$
$$+\beta E_{\boldsymbol{x'}}\left[V^m_{m,1}\left(c,x'_m,x'_f,\zeta\right)+V^m_{f,1}\left(c,x'_m,x'_f,\zeta\right)\right]-\left[V_{m,out}\left(c,x_m,c',\tilde{\xi}_m\right)+V_{f,out}\left(c,x_f,c',\tilde{\xi}_f\right)\right]. \tag{19}$$

It follows that the value for a married woman, conditional on receiving a mobility shock, is

$$V^m_{f,of}\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right)=V_{f,out}\left(c,x_f,c',\tilde{\xi}_f\right)$$
$$+\max\left\{0,\eta S_{stay}\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right),\eta S_{move}\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right)\right\}. \tag{20}$$

There are two policy functions associated to the latter equation. The first, $d_{of}\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right)$, captures the divorce choice of a couple and equals one if divorcing is optimal for a (x_m,x_f,ζ)-type couple living in c and holding the opportunity to move to c' with associated preferences $\tilde{\xi}_m$ and $\tilde{\xi}_f$. The second, $t\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right)$, captures the optimal migration choice conditional on not divorcing and equals one if migration is optimal and zero otherwise. Notice that, because of the assumption of linear utility and Nash bargaining there is always agreement on divorce between the involved parties.

Finally, we can write the Bellman equation for a married woman before the realization of the mobility shock as

$$V^m_{f,2}\left(c,x_m,x_f,\zeta\right)=(1-\chi)\,V^m_{f,nof}\left(c,x_m,x_f,\zeta\right)$$
$$+\chi\sum_{c'|c'\neq c}\theta\left(c'|c\right)\int_{\tilde{\xi}_m}\int_{\tilde{\xi}_f}V^m_{f,of}\left(c,x_m,x_f,\zeta,c',\tilde{\xi}_m,\tilde{\xi}_f\right)df_\xi\left(\tilde{\xi}_m\right)df_\xi\left(\tilde{\xi}_f\right) \tag{21}$$

3.3 Stationary Distribution

Let the total mass of men (married and single) in the economy be M and the total mass of women be F. Let $\tilde{\mu}_{c,x_m,x_f,\zeta}$ be the mass of married households of type (x_m,x_f,ζ) living in c. It must be that

$$\sum_c\int_{x_f}\left(d\mu_{c,f,x_f}+\int_{x_m,\zeta}d\tilde{\mu}_{c,x_m,x_f,\zeta}\right)=F \tag{22}$$

and similarly for men

$$\sum_c\int_{x_m}\left(d\mu_{c,m,x_m}+\int_{x_f,\zeta}d\tilde{\mu}_{c,x_m,x_f,\zeta}\right)=M. \tag{23}$$

To complete the definition of a stationary equilibrium, we need a set of equations equalizing the inflows and outflows for each type on single men and women. Given the length of such equations, they are reported in appendix D (equations (29) to (34)).

3.4 Equilibrium

In equilibrium, all agents act optimally with respect to their marriage and mobility choices taking the endogenous marriage market matching probabilities as given. The equilibrium is defined as a consistency requirement between the latter probabilities and the stationary distribution induced by the behavior of agents.

Definition. *A stationary equilibrium consists of*

i *a set of policy functions for migration of singles* $\left\{t_g\left(c, x_f, c', \tilde{\xi}_f\right), t_{out,g}\left(c, x_f, c', \tilde{\xi}_f\right)\right\}_{g=m,f}$ *and couples* $\left\{t\left(c, x_m, x_f, \zeta, c', \tilde{\xi}_m, \tilde{\xi}_f\right)\right\}$, *and policy functions for marriage formation* $\{m\left(c, x_m, x_f, \zeta\right)\}$ *and marriage dissolution* $\left\{d_{nof}\left(c, x_m, x_f, \zeta\right), d_{of}\left(c, x_m, x_f, \zeta, c', \tilde{\xi}_m, \tilde{\xi}_f\right)\right\}_{g=m,f}$,

ii *a set of value functions for singles* $\left\{V^s_{m,t}, V^s_{m,t}\right\}_{t=1,2}$ *and married households* $\left\{V^m_{m,t}, V^m_{f,t}\right\}_{t=1,2}$ *and (iii) a distribution of singles,* $\left\{\mu_{c,g,x_g}\right\}_{c,x_g}$*for* $g = m, f$, *and married households,* $\left\{\tilde{\mu}_{c,x_m,x_f,\zeta}\right\}_{c,x_m,x_f,\zeta}$,

such that

1. *Given the distribution of singles, which determines the matching probabilities in the marriage market, the value functions solve equations (7), (10), (12) and (21), with associated policy functions.*

2. *Given the policy functions, the distributions of single and married households solve equations (29) through (34) in appendix D.*

We now state the formal existence proposition.

Proposition (Existence). *Under regularity conditions, a stationary equilibrium exists.*

Proof. See appendix E. □

We do not provide a formal proof for the uniqueness of the equilibrium and, thus, our model is open to the existence of multiple equilibria. Nevertheless, in the practical application, we have found that, for a given set of parameters, the model always converges to the same stationary equilibrium.

4 Structural Estimation

In this section, we describe the strategy we implement to estimate the model. First, we describe our choices aimed at reducing the computational burden of solving and estimating the model. Secondly, we describe the estimation procedure and finally we discuss results and model fit.

4.1 Details on the Numerical Implementation

Numerically solving the model is an extremely burdensome exercise. Despite its simplicity, 209 cities and 95 occupations make the state space too big, rendering the numerical solution of the model unfeasible. In other words, the model suffers from the curse of dimensionality.

To address this problem, we reduce the number of cities. The choice of reducing the number of cities is driven by two considerations. First, since one of the goals of the paper is to assess the marital sorting pattern across different occupations, it seems natural to preserve as much occupational heterogeneity as possible. Secondly, the reduction of the number of cities is just as effective at reducing the computational burden as cutting the number of occupations. To see this, it is sufficient to notice that the number of operations to be carried out to solve the fixed point problem described by the Bellman equations is proportional to C^2J^2, where C is the number of cities and J the number of occupations.[35]

Having established the goal of reducing the number of cities, we are left with the quest of defining them. In the model, the only exogenous driver of the migration and marriage patterns is the wage variation across cities and occupations.[36] In particular, the geographic variation of wages within an occupation determines the potential gains from migration while the way in which the wages of occupations covary across cities shapes the marriage choices.[37] It follows that, for the model to do a good job at replicating the data, we need to preserve the aforementioned covariance structure as much as possible. A natural trade-off arises: the more cities we include the closer we can replicate the empirical covariance structure in the model but the more computationally costly it becomes to solve the model.

In practice, we divide (cluster) the 209 original cities into 25 groups such that cities with similar estimates of the city-occupation fixed effect $\alpha_{c,o}$ belong to the same group,

[35] The biggest object in the model is the policy function describing the migration decisions of couples. If P and L are the number of points used to discretize the support of the preference and bliss shocks respectively, its dimensionality is given by $C^2J^2P^4L$.

[36] The other driver is the differential in the marriage market conditions which is endogenous.

[37] Wage levels are relevant only to the extent that the returns to scale in the consumption of married agents, captured by $x(\cdot,\cdot)$, introduces a motive for assortative mating on income. With Nash bargaining and linear utility, in the absence of returns to scale in consumption, wage levels do not affect the marital surplus.

and treat each group as a single city. The clustering process involves two steps: first, we perform a Principal Component Analysis on the city-occupation premia and, then, we apply an unsupervised machine learning clustering algorithm (k-means) to a subset of the principal components to obtain 25 clusters. The number of clusters has been chosen as a compromise between the need for speed and the necessity to preserve the wage covariance structure. The details of this procedure and the resulting partition of MSAs are reported in appendix G.

In the model, wages are exogenous and heterogeneous with respect to gender and occupations. In reality, other factors also affect wages (e.g., education). Moreover, these additional sources of heterogeneity can induce additional marital sorting directly and not only through income. For instance, education constitutes both a source of marital sorting in itself and through income, due to the college premium. Ideally, one would like to explicitly model all the additional sources of wage heterogeneity but, in this setting, it is not a computationally viable option. Nevertheless, we can still partially account for the effects of these factors that come through wages. To do so, we compute the wage distribution to be fed to the model taking into account the demographic composition of each occupation. First, we estimate equation (1) with the fictitious cities, then we divide the sample by occupation and, for each subgroup, we compute the predicted wage for each city assuming all workers are men (i.e., as if the whole subgroup were to live and work in the same city). The model counterpart of each wage is then obtained as the average annualized predicted wages for each city and occupation.[38] Women's wages are obtained by simply subtracting the estimated wage gap (19.3%).

Job transitions are also exogenous to the model. We compute two gender-specific transition matrices for occupations from yearly CPS-ASEC data from 2018 to 2017 applying the same sample selection restriction as in the ACS samples. The initial occupation is drawn from the stationary distribution of the Markov chain corresponding to the relevant transition matrix. As shown in figure 10 in appendix B, this distribution is very close to the distribution of occupations obtained from the ACS sample.

Finally, we approximate both the marriage quality and city preference processes by a three-state Markov process obtained through the Rouwenhorst (1995) method.

Despite all the simplifications, the model still requires substantial computational power. For our calculations, we relied on the resources for high-performance computing provided by the Digital Research Alliance of Canada.

[38]The predicted wages are measured as the log of hourly wages. We take the exponential of the average predicted value and multiply it by 2080 (40 hours per week over 52 work weeks).

Parameter	Meaning	Value	Source
η	Bargaining Power	0.5	Equal bargaining power for spouses
α	Curvature of matching function	0.5	No gender bias in matching probabilities
λ	Overall matching efficiency	0.151	Goussé, Jacquemet and Robin (2017)
$\tilde{\beta}$	Time preference	0.98	Attanasio, Low and Sánchez-Marcos (2008)
ρ	Economies of scale in consumption	0.777	McClements scale
$1-\omega$	Death probability	1/30	A 30-year lifespan
ρ_ζ	Persistence of bliss shock	0.959	Greenwood et al. (2016)

Table 6: Preset parameters.

4.2 Estimation

There are 15 parameters to be chosen but not all the parameters can be directly identified from the data. Seven of these are set to values drawn from the literature or to reasonable values. The rest of the parameters are estimated.

Preset Parameters

The list of preset parameters and their values are reported in table 6.

Each period in the model corresponds to a calendar year. To conform with the empirical analysis, we assume that agents enter this economy at age 25 and they expect to remain in the labor and marriage markets for 30 years on average (this determines ω). We calibrate the economies to scale in consumption for married households ρ to match the McClements, according to which for the same level of consumption a person living alone spends 61% of what a childless couple spends.[39] The matching function is calibrated to be unbiased ($\alpha = 0.5$) and the scale parameter λ if taken from Goussé, Jacquemet and Robin (2017).[40] The bargaining power parameter η is set to 0.5, such that both spouses have the same bargaining power. Finally, the time preference parameter is set to 0.98, taken from Attanasio, Low and Sánchez-Marcos (2008), while the persistence of the marriage quality ($\rho_\zeta = 0.959$) comes from Greenwood et al. (2016).

Indirect Inference

We estimate the remaining parameters, the vector $\mathbf{\Pi} = \left(\chi\ \kappa\ \delta\ \gamma\ \rho_\xi\ \sigma_\xi\ \bar{\zeta}\ \sigma_\zeta\right)'$, by indirect inference (Gourieroux, Monfort and Renault, 1993). We estimate a set of auxiliary parameters from real ($\hat{\boldsymbol{\phi}}_{\mathbf{data}}$) and simulated data ($\boldsymbol{\phi}_{\mathbf{sim}}$) and choose $\mathbf{\Pi}$ to minimize the Euclidean

[39] In this model, the total expenditures of a married household equals the sum of the wages, $x = w^M + w^F$. Under the assumption that both spouses have the same consumption c and the same wage w we have that $2c = (w^\rho + w^\rho)^{\frac{1}{\rho}} = 2^{\frac{1}{\rho}} w = 2^{\frac{1}{\rho}} \frac{x}{2}$ which, using $c = 0.61x$, gives $\rho = \frac{\log(2)}{2\log(2)+\log(0.61)} = 0.777$.

[40] They also use a Cobb-Douglas specification for the matching function in their model.

distance between them. The estimated vector $\hat{\mathbf{\Pi}}$ is given by

$$\hat{\mathbf{\Pi}} = \arg\min_{\mathbf{\Pi}} \left(\hat{\boldsymbol{\phi}}_{\mathbf{data}} - \boldsymbol{\phi}_{\mathbf{sim}}(\mathbf{\Pi})\right)' \left(\hat{\boldsymbol{\phi}}_{\mathbf{data}} - \boldsymbol{\phi}_{\mathbf{sim}}(\mathbf{\Pi})\right). \tag{24}$$

The auxiliary model is composed by 18 auxiliary parameters (targets).

The first set of targets relates to marriage market outcomes and includes the average marriage rate of men, the average divorce rate of men, and the fraction of new marriages that are occupationally homogeneous. Moreover, it also includes all the estimated coefficient on the geographic mismatch dummies from the estimation of the marriage and divorce probability model for men.[41]

The second set of targets relates to migration and includes the migration rate of couples and single men and the set of average changes in geographic mismatch conditional on migrating.

Identification

As it is usually the case with structural models, the identification of structural parameters is jointly determined by several variations in the data. Nevertheless, it is possible in this model to give a rather clear, partial equilibrium intuition of how every single parameter is estimated from the data.

The probability of getting a mobility shock, χ, is identified by the average moving rate. Similarly, the cost of divorce, δ, and the average of the bliss process, $\bar{\zeta}$, are identified by the average divorce rate and the average marriage rate respectively. The parameter γ, which determines the extent to which occupational homogamous matches are more likely to be formed in the marriage market, is identified by the fraction of newlyweds who work in the same occupation.

The standard deviation of the preference shock, σ_ξ, is identified by the differences in the average change in geographic mismatch conditional on migrating. The standard deviation of the preference shocks determines the relative importance of preferences and wage distributions in determining the migration patterns. If it is high enough, agents will migrate based solely on their preferences and wages would play a little role. In this case, households would move randomly and the average change in geographic mismatch would reflect this, with the best geographically matched households dropping, on average, by 2.5 categories, and badly matched households gaining as much. Conversely, as this standard deviation falls towards

[41] Given the absence of heterogeneity in education in the model, we estimate these two probability models without the education dummy. The results are reported in the first column of tables 18 and 22 in appendix B.

zero, wage differentials become relatively more important causing households to move on average more and more towards high-paying cities. This implies that the average change in geographic mismatch increases for all households, but more for those already living in good cities (who would only rarely move and only among the best cities) than for the others (who still find profitable to move almost anywhere).

The standard deviation of the bliss shock, σ_ζ, is identified by the differences in the marriage probabilities across individuals with different levels of mismatch. The logic behind the identification is similar to the previous one. For a high standard deviation of the bliss shock, the impact of mobility considerations on marriage decisions is dwarfed by the size of the shock. In this case, the model will produce no differences in the marriage rates across levels of mismatch. As we reduce it, mobility considerations become more and more important and the model will start generating a higher differential in the marriage hazard rates between well- and badly matched households.

The persistence of the preference shock, ρ_ξ, is similarly identified by the differences in the divorce rates. On the one hand, given the high value of the migration option, marriages involving badly matched spouses are characterized by lower surpluses and are thus less stable. Moreover, because of the endogeneity marriage, spouses in such matches tend to have higher-than-average preference shocks for the current location. A lower persistence of the preference shock means a faster reversion to the mean and thus a faster fall in the marital surplus for these couples, which makes divorce more likely. On the other hand, marriages involving well-matched individuals can be profitable even for lower-than-average preferences for the current city. The faster mean reversion will then make migration-induced divorces less likely. In other words, increasing the persistence of the preference shock reduces the difference in the probability of divorce across households.

Finally, the migration cost, κ, is identified by the difference in the mobility rates of single men and couples. Upon receiving a mobility shock, the migration choice of a single individual follows a threshold rule according to which migration occurs only if the draw from the preference distribution is large enough. Married households follow a similar threshold rule but, in this case, the relevant distribution is not the preference distribution itself but the distribution coming from the sum of two preference parameters, since each spouse draws a preference parameter independently. The latter distribution is then more dispersed the former. This implies that, the probability of moving for single households will increase relatively more after a marginal reduction of the moving cost κ than the probability of moving for married households which gives the identification.

Figure 25 in appendix B illustrates identification by showing how the identifying variables change as we modify the value of each estimated parameter.

Parameter	Meaning	Value
χ	Probability of a mobility shock	0.07
κ	Cost of moving	0.51
δ	Cost of divorcing	0.27
γ	Matching function bias	2.07
ρ_ξ	Persistence of preference shock	0.03
Parameter	Meaning	Value ('000 of dollars)
σ_ξ	Std. Dev. of preference shock	64.32
$\bar{\zeta}$	Mean of bliss shock	-15.73
σ_ζ	Std. Dev. of bliss shock	0.85

Table 7: Estimated parameters.

4.3 Results

Parameters and Moments Fit

Table 7 lists the estimated parameters. The labor market friction parameter, χ, is estimated at 0.07 which corresponds to each household getting a mobility shock roughly once every 15 years. The low value of χ is necessary to match the low migration rate, especially because of the lack of lifecycle effects that would otherwise naturally reduce the migration rates without the necessity of strong frictions. The cost of moving is estimated to be 51% of yearly income. In the stationary distribution, for the average mover, the cost of migration is \$26,651.[42] The cost of divorce is estimated to be 27% of annual income. For the average divorcee, this cost amounts to \$11,399.[43] The matching function bias parameter γ is estimated to be 2.07. In a completely homogeneous and uniform marriage market in which an equal number of single men and women are evenly spread across occupations, this value implies that the probability for any agent of meeting a potential spouse working in the same occupation conditional on any match occurring is 3.16%.[44] The estimated mean value of the bliss shock, ζ, is -\$15,732. Negative values for such a quantity are not uncommon in the literature.[45] A negative mean for the bliss shock is mechanically needed to counteract the gains from economies of scale.

[42] Our estimate is much smaller than the estimate of Kennan and Walker (2011) (above \$300k) but bigger than that of Gemici (2016) (\$5-10k). This is not surprising given the important differences across the three models. The former estimates a frictionless model with exogenous income which requires huge moving costs to match the fairly low observed moving rates. In the latter, migration is frictionless but the model comprises a psychic cost from relocation which is increasing in the time spent in one location which contributes to lowering the monetary cost of migration needed to match the empirical moments.

[43] A real-world benchmark for the monetary cost of divorce, mentioned also by Voena (2015), is provided by the Rosen law firm fee calculator. According to it, the legal fees associated with a divorce by agreement (the most common kind) range between \$6,000 (for marriages involving no children and few assets), and \$38,000.

[44] In comparison, this value is 1.05% if $\gamma = 0$.

[45] See Greenwood et al. (2016).

Moment	Data	Model
Average probability of migration (single men)	0.76%	0.93%
Average probability of migration (couples)	0.24%	0.10%
Average probability of marriage (men)	8.35%	9.12%
Average probability of divorce (men)	2.69%	0.56%
Fraction of newlyweds in the same occupation	8.54%	8.28%

Table 8: Other targeted moments.

Without the latter, the model would require a positive value for the bliss parameter to make marriages profitable.

Figure 9 shows the auxiliary parameters concerning the relation between geographic mismatch and migration, marriage and divorce, and compares the values estimated from real and simulated data. Table 8 shows the remaining targeted moments.[46] Many of the empirical moments are closely reproduced by the model. Nevertheless, the model underestimates the average migration rate of couples which is less than half its empirical counterpart. One possible justification for this discrepancy is the absence of lifecycle effects in the model which make it impossible for the model to produce migration flows linked, for instance, to the prospect of retirement. The model also underestimates the divorce rate. The estimated rate is, in fact, about a fifth of its empirical counterpart. This discrepancy can be partly justified by the fact that the model lacks many possible sources of marital instability. The most obvious is the lack of employment dynamics and, more in general, any income dynamics besides those implied by migration, which has been shown to impact marital stability (Swensen, Lindo and Regmi, 2020).

Model Fit: Non-Targeted Moments

The estimation strategy does not make use of any moment obtained from the geographic distribution of workers. This makes the latter the optimal candidate to test the performance of the model. In figure 10, we compare a measure of the geographic concentration of employment for each occupation, the Herfindahl–Hirschman (HH) index, as obtained from the data and the model. For clarity, both axes are displayed on a logarithmic scale.[47] Each bubble is an occupation and the bubble size is proportional to the total employment in the given occupation.[48] Overall, figure 10 shows that, despite its simplicity, the model does a good job at matching the geographic concentration of employment.

[46] In appendix B, figure 26 and table 24 show the equivalent of the targeted moments for women.

[47] The same graph in absolute scale is shown in figure 27 in appendix B.

[48] We compute total employment from the data. The employment structure in the model mirrors the empirical one. See the discussion in appendix B and figure 24.

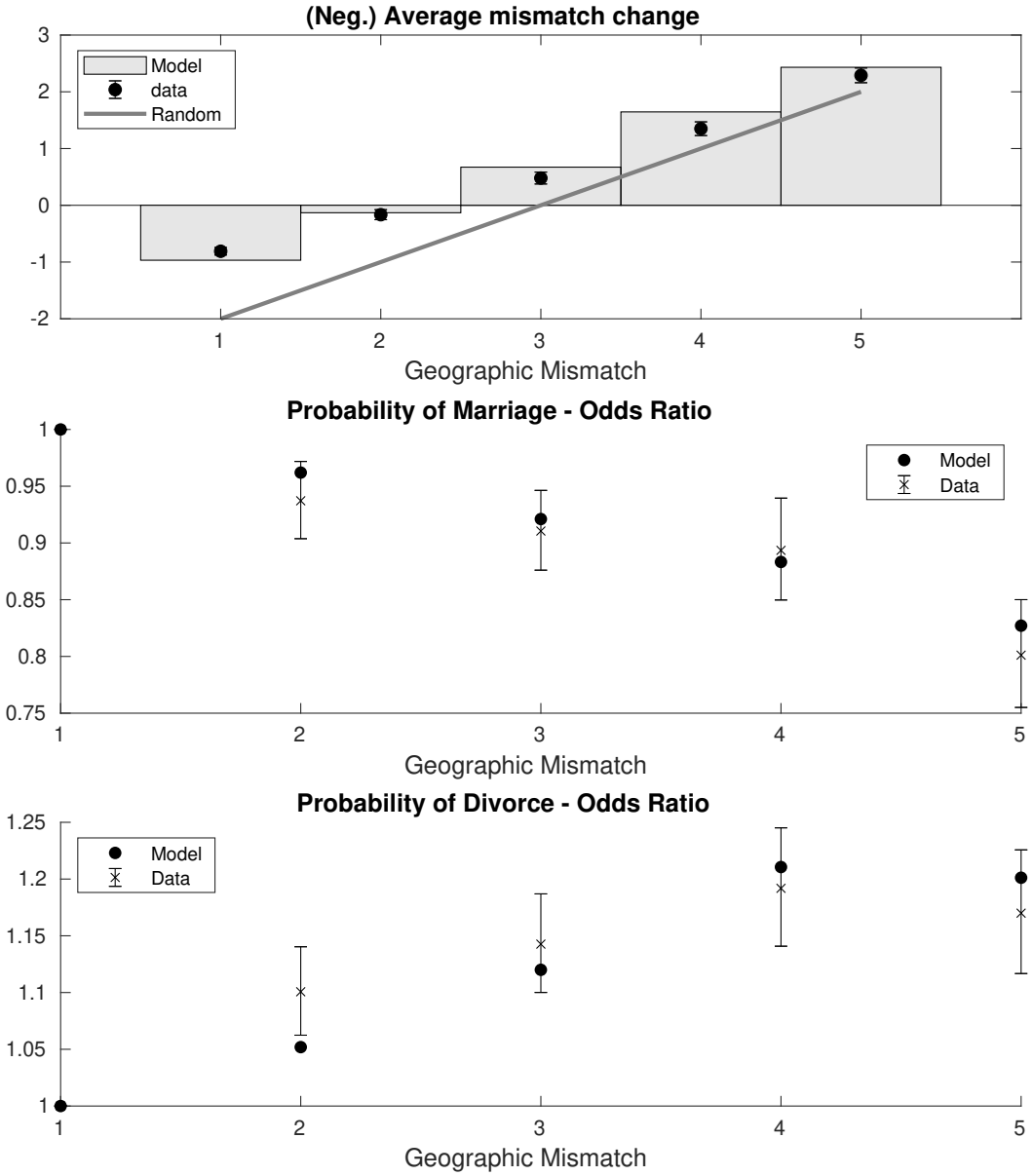

Figure 9: Estimates of the auxiliary model from real and simulated data.

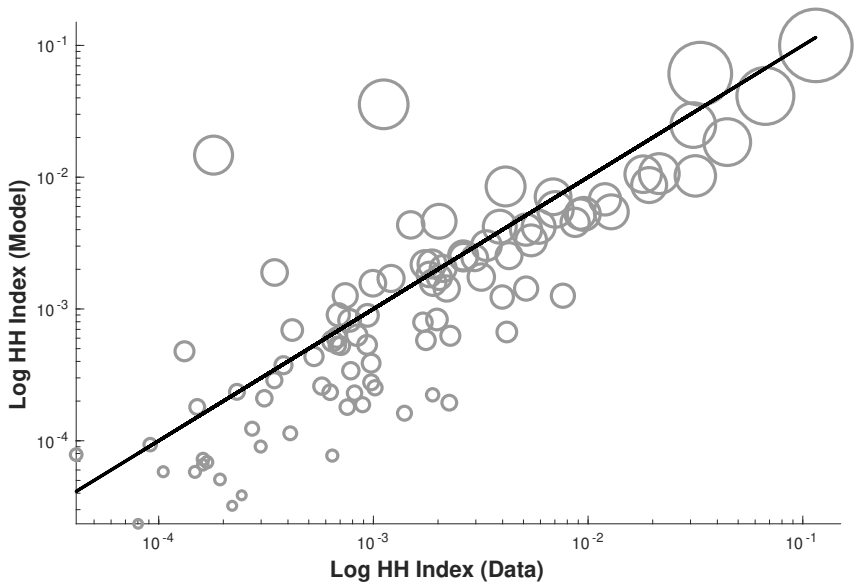

Figure 10: Goodness of fit: HH index of the geographic concentration of employment by occupation (log-log scale).

In addition, figure 11 shows the simulated version of our preferred measure of assortative mating (eq. 4), namely the simulated counterpart of the bottom quadrants of figure 5.The right panel of the figure also reports the empirical values for comparison. The model cannot perfectly match the data, but the figure shows that it can reproduce all the qualitative features.[49] This is not surprising since the empirical measure embeds other sources of assortative mating that might correlate with geographic mismatch but that are not present in the model.

Finally, the model is also able to successfully reproduce the ex-post gender wage gap: 26.1% in the data and 26.2% in the model.[50]

5 Counterfactual Experiments

In this section, we use the estimated model to perform a few counterfactual exercises to study the effects of the interaction between migration and marital choices. First, we analyze

[49]The drop in the measure of assortative mating from mismatch level 4 to level 5 is caused by the fact that highly mismatched individuals tend to have a higher preference for the city they live in. In such a marriage market, in which many people wouldn't move even if given the opportunity, the motif for assortative mating on geographic mismatch is weaker. The drop in the measure of assortative mating is present in the data as well even though the magnitude is much smaller. This difference might be due to the presence of additional sources of assortative mating that correlate with geographic mismatch but also to the coarseness of the approximation used to model the processes that describe the evolution of preference shocks and the bliss parameter, both of which are approximated by 3-state Markov chains.

[50]See appendix H.

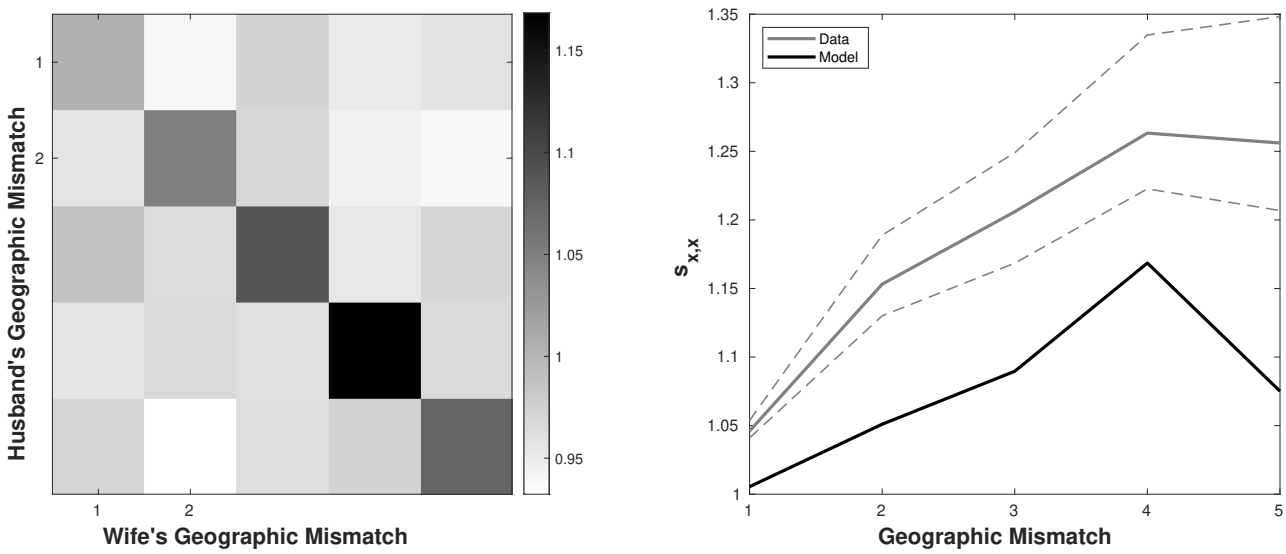

Figure 11: Model-generated assortative mating (corresponding to equation 4). The left panel shows all the values while the right panel focuses on the diagonal elements and compares them to the empirical values.

the cost of family ties in terms of forgone income growth. Secondly, we show that marriage improves the geographic allocation of the labor force by increasing the concentration of workers in more productive areas and we study the implications of this for income inequality. In this regard, we also show, through a simple tax experiment, that incentivizing marriage can favor the concentration of the labor force in more productive areas. Finally, we look at how migration considerations affect marriage market outcomes.

5.1 The Impact of Marriage on Labor Market Outcomes

Marriage reduces geographic mobility reducing the capacity of married workers to take advantage of otherwise profitable job opportunities that require a geographic relocation. To what extent do family ties reduce the wage income growth of workers? To answer this question, we compare the migration behavior and income path of simulated individuals who marry at least once over 35 years (the time span considered in the empirical analysis) to the same outcomes for the same individuals in the counterfactual scenario where they never marry.[51]

Table 9 shows the fraction of workers who move at least once in the baseline and counterfactual simulation, conditional on marital history. In the baseline simulation, 9.1% (9.2%)

[51] To isolate the impact of family ties, in the counterfactual calculation we keep all the migration incentives coming from the model alive. This includes the incentive of migration linked to marriage market conditions. In other words, in the counterfactual world, individuals still consider better marriage markets as an amenity when deciding to move. The difference with the baseline model is that marriage never actually realizes.

		Men	Women
Married at least once	Baseline	9.1%	9.2%
	Counterfactual	16.9%	17.3%
Married at 25	Baseline	4.0%	4.2%
	Counterfactual	12.7%	13.1%
Never Married		12.2%	11.7%

Table 9: Fraction of individuals who migrated at least once over 30 years.

of men (women) who married at least once also move at least once over 35 years. Without the marital bond, this fraction rises to 16.9% (17.3%). This increase in mobility is even more striking for those who, in the baseline simulation, marry at age 25. In the baseline, only 4.0% (4.2%) of men (women) end up moving at least once, while, in the counterfactual scenario, this fraction is roughly tripled, reaching 12.7% (13.1%). The table also reports the fraction of migrating workers who, in the baseline specification, never marry. This fraction is above the baseline for those marrying at least once, but below the counterfactual level.

In figure 12, we show the average ratio of yearly income in the counterfactual over the baseline simulation for individuals who marry at age 25 as a function of their initial geographic mismatch. Not surprisingly, the more severe is the initial geographic mismatch in the labor market, the more costly is marriage in terms of lost income growth. Men who are highly geographically mismatched (level 5) at age 25 experience, in the counterfactual scenario, a higher growth of yearly labor income, ending up earning about 7.6% more by age 35 and about 10.1% more by age 55. For women, the same figure is 4.9% by age 35, and 8.2% by age 55. The extent of the gains is lowered for men (women) who are initially better matched with their local labor market, with those with mismatch level 4 gaining about 3.0% (2.4%), and those with levels 3 and 2 about 1.9% (1.2%) and 0.6% (-0.5%) respectively. For the initially well matched, there is, in terms of labor income, a slight gain from entering an early marriage (-0.2% and -0.6% for men and women, respectively). This is due to the fact that, because of preference shocks, some married individuals in high-paying cities would receive higher utility relocating to lower-paying cities, but marriage prevents them from doing so. Despite being already sizable, these numbers should be interpreted as lower bounds to the actual cost of marriage as the model does not feature within-city income growth, which is substantially different across geographic areas (Baum-Snow and Pavan, 2012).

5.2 Marriage and the Geographic Distribution of Labor

In this section, we use our model to answer the following question: what is the effect of marriage and marriage markets on the geographic allocation of labor? With the previous

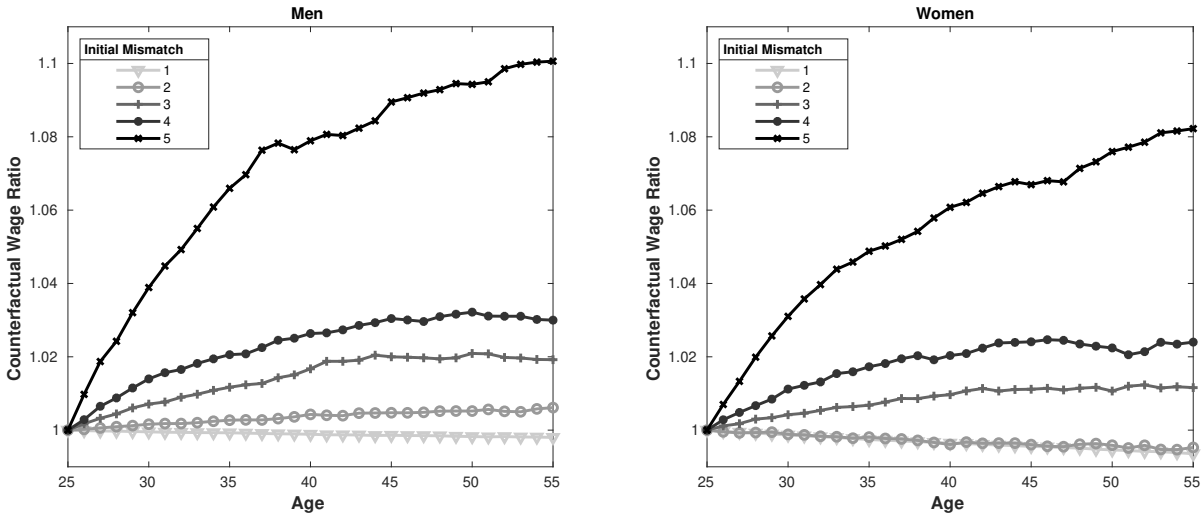

Figure 12: Average ratio of yearly income in the counterfactual vis à vis the baseline simulation, for individuals who marry at age 25, as a function of their initial level of geographic mismatch.

exercise, we have shown that, by entering an early marriage, individuals give up migration opportunities that lead to wage growth. At the stationary equilibrium, an early marriage comes with an implicit idiosyncratic cost in terms of lost wage growth. This suggests that, by reducing mobility, marriage causes a misallocation of labor preventing some individuals from relocating to more productive cities. Nevertheless, while it is true that marriage restricts mobility, it is also true that marriage, and in particular the endogenous heterogeneity of marriage market conditions, constitutes a motive for migration. Conditional on one's occupation, high-paying cities feature a higher concentration of single individuals working in such an occupation. Since singles working in the same occupation are more likely to marry each other, high-paying cities present better marriage markets. This constitutes an additional incentive for geographically mismatched individuals to relocate to better cities.

To quantify the effect of marriage on the geographic distribution of the labor force, we compute a counterfactual equilibrium in which marriage markets are completely shut down (Counterfactual 1).[52] The left panel of figure 13 reports the sizes of the modeled cities in the baseline and counterfactual scenarios. The plot shows that, without marriage markets, the population in the economy would be much less concentrated, with 3 of the biggest cities shrinking in size by 32% on average, while all the other cities become 67% bigger on average.

[52]There is an essential difference between this experiment and the one performed in the previous section. In the latter, the counterfactual scenario was computed by simply setting the marriage policy functions to zero while here we recompute the whole equilibrium shutting down the marriage market completely. This implies that in the previous case migration decisions also incorporate incentives coming from differentials in marriage market conditions across cities, while here this is not the case.

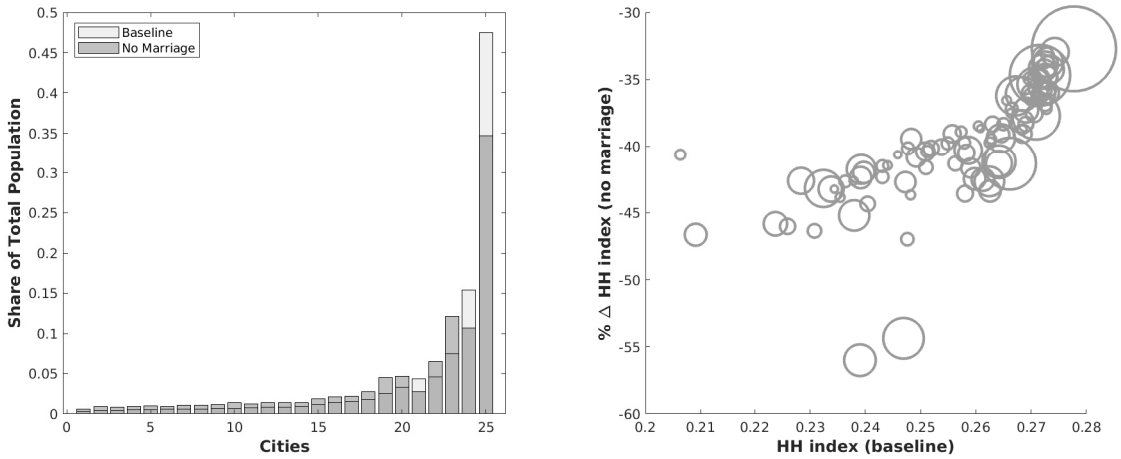

Figure 13: The effects of marriage on the geographic distribution of workers. Left: fraction of the total population residing in each city. Right: change in the geographic concentration (HH index) of employment; each bubble is an occupation and the size of the bubble is proportional to the total employment in such occupation.

The right panel of the figure looks at the changes in the geographic concentration of each occupation, as measured by the HH index, as a function of the initial concentration. All occupations see their concentration decrease by at least 30%.

A back-of-the-envelope calculation allows us to get a sense of how the presence of marriage markets affects the aggregate wage bill in the baseline equilibrium relative to the one in the counterfactual scenario with no marriage markets. As shown in the second column of table 10, the presence of marriage markets causes an increase in total labor earnings of 2.70% over the counterfactual scenario. Since our model does not allow for wage adjustments, we compute boundaries for this figure using estimated wage elasticities from the literature. The lower bound is computed using occupation-specific elasticities from Alonzo and Gallipoli (2020), who estimate a standard CES production function with heterogeneous labor inputs, while the upper bound is obtained using an estimate of the elasticity of wages to city size which captures agglomeration economies from the meta-analysis by Melo, Graham and Noland (2009).[53]

The estimated change in total earnings is due to the combined effect of two factors. On the one hand, introducing marriage in a world without it provides an additional incentive for singles to migrate to more productive cities to take advantage of better marriage market conditions. In these cities, acceptable partners are on average richer and the endogenous migration flows improve marriage market conditions even further by increasing the share of acceptable partners. On the other hand, by restricting mobility, marriage prevents

[53]For details see appendix I.

	Counterfactual 2	Baseline
Changes in total labor earnings over Counterfactual 1	+1.09%	+2.70%
	(0.00 - 2.27)	(1.66 - 2.85)

Table 10: Total labor earnings changes relative to the counterfactual scenario with no marriage markets. The first column refers to the counterfactual scenario in which individuals internalize the marriage market conditions but marriage never actually realizes. The second column refers to the baseline equilibrium.

the geographic reallocation of labor. To disentangle the two effects, we compute a second counterfactual scenario in which agents take into account the possibility of marrying when evaluating migration, with the matching probabilities taken from the baseline equilibrium, but in which marriage never actually happens.[54]

The first column of table 10 shows the change in total labor earnings relative to the counterfactual equilibrium with no marriage market. The additional incentive to relocate to more productive cities, provided by the different marriage market conditions, accounts for about 40% of the total gain in aggregate earnings generated by the presence of marriage markets. This suggests that the remaining 60% is caused by the fact that marriage restricts migration. That marital ties to mobility have a positive effect on total labor earnings is perhaps surprising at first. Nevertheless, there is a simple explanation. As shown in the previous section, marital ties do indeed prevent geographically mismatched individuals from relocating to more productive cities, but it also prevents the well-matched to move from high-paying cities to low-paying cities if their idiosyncratic preferences dictated so. Since in equilibrium the second group largely outnumbers the first, and this group is more likely to enter marriage and less likely to exit it through a divorce, the positive effect of preventing well-matched individuals from relocating to low-productivity cities is bigger than the negative effect of preventing mismatched individuals to move to better cities.

By affecting the geographic distribution of the labor force, marriage also has implications for labor income inequality. High inequality, both at the national and at the local level, has been associated with a variety of outcomes. Persson and Tabellini (1994) and Alesina and Rodrik (1994) argues that in democracies, high inequality at the national level spurs redistributive policies that end up harming growth. The same negative relationship between income inequality and growth has been found at the local level by Glaeser, Resseger and Tobio (2009). Fajnzylber, Lederman and Loayza (2002) and Daly, Wilson, and Shawn (2001) find a positive causal relationship between inequality and violent crimes. Figure 14 shows

[54]In practice, we compute the value functions and the associated policy functions assuming that the matching probabilities are the same as in the baseline equilibrium. Given these policy functions, we compute the stationary equilibrium assuming there are no matches in the marriage market.

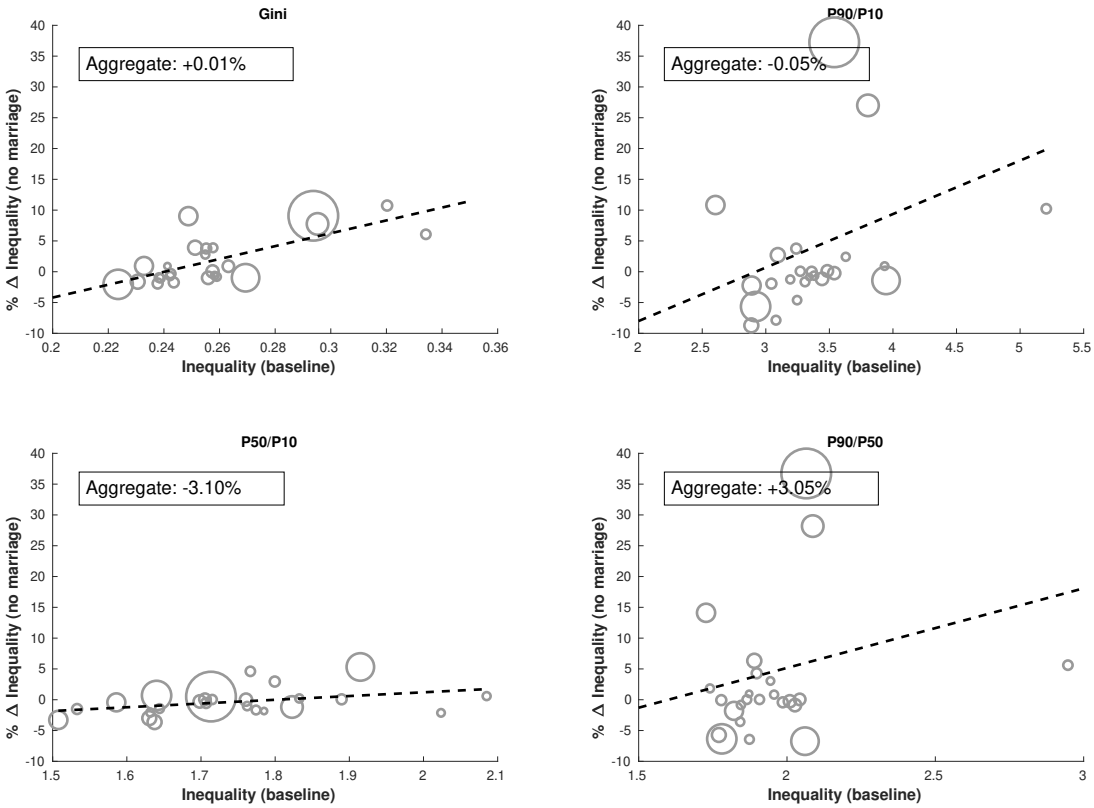

Figure 14: Changes in labor income inequality as measured by four different indexes. Each bubble is an occupation and the size of the bubble is proportional to the total employment in such occupation. The text box on the top-left corner of each panel reports the change in aggregate inequality as measured by each index.

the changes in the Gini coefficient and several percentile ratios for each city when moving from the baseline model to the counterfactual without marriage. In each panel, a bubble corresponds to a city and its size is proportional to the city's population. The text box on the top-left corner of each panel also reports the change in aggregate inequality as measured by the corresponding inequality index. We observe a heterogeneous effect of marriage on income inequality. Regardless of the specific measure, we see a positive relationship between the inequality level and its change. Moreover, we see that the migration patterns induced by marriage cause unequal cities to become more equitable and vice versa. A closer look at the magnitudes of the changes in the percentile ratios reveals that most of the changes in the within-city inequality occur in the right tail of the income distribution. This asymmetry is also present in the aggregate where, while the Gini coefficient and the P90/P10 ratios change by little, the other two ratios show that marriage increases inequality among the poor and decreases it among the rich.

5.2.1 A Tax Experiment

In our model, a world without marriage is equivalent to a world where marriage is so invaluable that it is never optimal. From this perspective, the analysis of the previous counterfactual scenario suggests that anything that can affect the value of marriage, such as changes in social norms, technology, or policies, can also have implications for the geographic distribution of the labor force, with consequences for productivity and welfare. Here, we present a simple tax experiment to illustrate how policies affecting the value of marriage can be welfare improving.

The experiment consists of incentivizing marriage through an income tax break for married individuals, financed by taxing the income of singles. Let τ^M and τ^S be, respectively, the tax break for married individuals and the tax applied to singles, such that after the tax adjustments the income of married individuals ($w^M_{g,c,j}$) and that of singles ($w^S_{g,c,j}$) are given by

$$w^M_{g,c,j} = (1+\tau^M) w_{g,c,j} \tag{25}$$

$$w^S_{g,c,j} = (1-\tau^S) w_{g,c,j} \tag{26}$$

The government budget constraint is given by

$$\tau^S \sum_c \left(\int_{x_m} w_{c,m,j_m} d\mu_{c,m,x_m} + \int_{x_f} w_{c,f,j_f} d\mu_{c,f,x_f} \right) = \tau^M \sum_c \left(\int_{x_m,x_f,\zeta} \left(w_{c,m,j_m} + w_{c,f,j_f} \right) d\tilde{\mu}_{c,x_m,x_f,\zeta} \right) \tag{27}$$

Assuming $\tau^M = 0.02$, the balanced budget condition requires $\tau^S = 0.08$. This is essentially introducing an income tax wedge between married and single individuals of about 10 percentage points.

In a similar fashion to figure 13, figure 15 shows the effects of the tax on the geographic distribution of the labor force. The left panel reports the sizes of the 25 cities in the baseline model against the model with the tax policy and shows that the tax policy, making marriage relatively more attractive, has the effect of increasing the size of 3 of the biggest cities, causing the rest of them to shrink. The right panel shows the change in the geographic concentration of employment in each occupation. As for the introduction of marriage in a world without it, the tax change has the effect of increasing the geographic concentration of employment in all occupations. A back-of-the-envelope calculation, similar to the one performed in the previous section, suggests that this corresponds to an increase in total

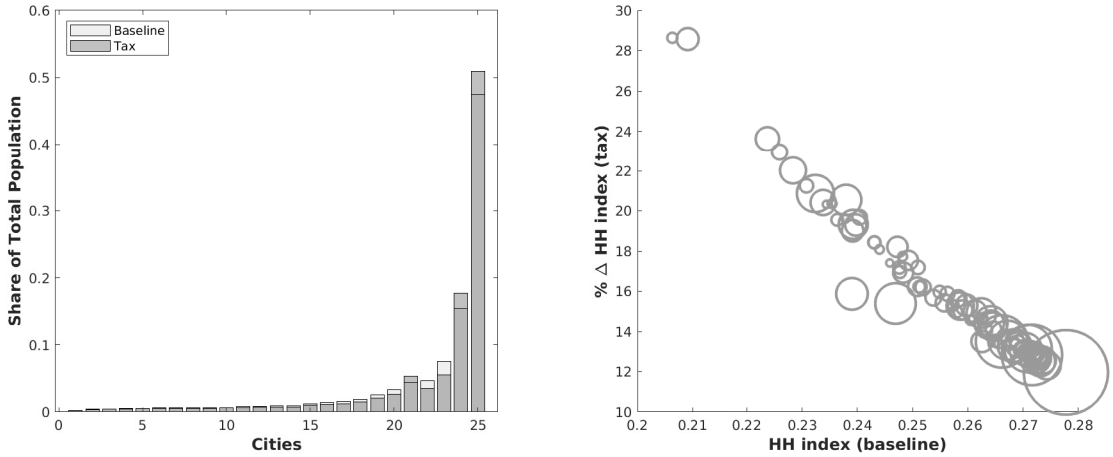

Figure 15: The effects of the tax policy on the geographic distribution of workers. Left: fraction of the total population residing in each city. Right: change in the geographic concentration (HH index) of employment; each bubble is an occupation and the size of the bubble is proportional to the total employment in such occupation.

labor earnings of 0.73% (0.69% - 1.16%). Moreover, due to the tax policy, the fraction of the population that is married increases from 69% to 78% and average welfare by 2.05%.

As before this is the result of two forces. On the one hand, the tax break makes marriage more attractive. Due to the proportionality of the tax break, the effect is stronger in the more productive cities that pay on average higher wages. This asymmetric effect makes marriage markets in these cities even more attractive to singles, pushing them to migrate there. On the other hand, marriage is now more valuable and, thus, more stable. This implies that it is also more constraining in terms of geographic mobility, causing geographically mismatched individuals to be less likely to marry relative to the geographically well-matched. As a consequence, fewer geographically mismatched individuals will be prevented from moving because of marriage while more geographically well-matched will not be able to move to less productive cities due to marital constraints.

5.3 The Impact of Migration on Marriage Market Outcomes

In this section, we analyze the counterfactual scenario in which agents are not allowed to migrate. Comparing the outcomes of the counterfactual marriage market to the baseline, we can isolate the impact of migration considerations on the marriage choices of individuals. Starting from the equilibrium distribution, we first simulate the counterfactual scenario in partial equilibrium, assuming that matching probabilities are fixed at the baseline values (we improperly call this the short-run and label it as "SR") and, secondly, we compute the new

	Marriage Rates				
	Baseline (%)	SR (%)	% Change	LR (%)	% Change
			All		
	8.19	8.47	+3.36	8.49	+3.60
Geographic Mismatch			Men		
1	9.33	9.50	+1.82	9.46	+1.41
2	6.96	7.39	+6.18	7.46	+7.14
3	6.07	6.61	+8.82	6.71	+10.55
4	6.74	7.15	+6.14	7.45	+10.63
5	4.46	4.74	+6.46	5.08	+13.98
Geographic Mismatch			Women		
1	9.21	9.40	+2.02	9.35	+1.49
2	6.98	7.44	+6.60	7.49	+7.29
3	5.76	6.28	+8.99	6.38	+10.75
4	6.68	7.10	+6.23	7.44	+11.31
5	4.20	4.50	+7.20	4.81	+14.44

Table 11: Marriage rates in the baseline and counterfactual scenario where migration is not allowed.

equilibrium (labeled "LR" for "long-run").

Table 11 shows the marriage rates for men and women and reports the percentage change of the counterfactual relative to the baseline. In what we call the short-run, the overall marriage rate increases by about 3.36%.[55] For both men and women, there is substantial heterogeneity across geographic mismatch levels with the change in the marriage rate ranging from +1.82% to +8.99%. The increase in the overall marriage rates is intuitively driven by the fact that, without the tensions linked to the possibility of a migration shock, many initially unprofitable marriages become suddenly profitable.

As we allow the marriage market to adjust to the new equilibrium (i.e., the long-run scenario), the removal of migration creates induces an increase in the marriage rates across all categories, with an overall increase of 3.6%. Moreover, there is a clear pattern in which the observed relative change in marriage the average marriage rate increases with the degree of geographic mismatch.

Perhaps surprising are the changes in the divorce rate shown in table 12. In the short run, we observe a 27.33% increase in the overall divorce rate. This increase in the divorce rate can be explained by the compounded effect of two factors. The first has to do with the fact that, in the baseline model, the presence of migration induces a form of marital sorting

[55]Notice that the overall marriage rate is the same for men and women by construction. The average marriage rate differs slightly from the empirical one since here we consider the overall marriage rate and not the rate of first marriage as we have done previously.

on geographic preferences. Without the possibility of migration, this form of marital sorting is removed and many of the marriages formed in the baseline model cease to be sustainable. In other words, in the baseline model, an important part of the marital surplus of existing marriages comes from the sorting on preferences induced by the possibility of receiving a mobility shock. Once mobility shocks are removed, this part of the marital surplus is lost as well which reduces the value of existing marriages. This mechanism is particularly important for geographically mismatched individuals as captured by the fact that the change in divorce rates increases with the level of geographic mismatch. Being relatively poorer, geographically mismatched individuals cannot resort to transfers to compensate spouses for living in cities they do not like much and, as a consequence, sorting on preferences is more important for them. The second factor is the increased marriage rate which makes the value of singlehood, the outside option in a marriage, more valuable. This implies that the marital surplus of existing marriages is reduced even more, prompting a further increase in divorce rates.

In the new equilibrium, the long-run scenario, the overall divorce rate is 0.40% bigger than in the baseline equilibrium. Nevertheless, there is substantial heterogeneity across men and women at different levels of geographic mismatch. These patterns are generated by two opposing forces. On the one hand, just like for the short-run, since marriage rates are now higher the average surplus from marriage is smaller which makes marriages more likely to be broken by a shock to the marriage quality (bliss) parameter. On the other hand, since the preference for the current city does not enter the marital surplus and mobility shocks do not happen, there are fewer shocks that might alter the value of marriage and cause a divorce.

Finally, the absence of migration affects the incentive to enter occupationally homogeneous marriages. As shown in table 13, both in the short and the long run we observe an overall drop in the fraction of homogeneous marriages of -2.37% and -2.90% respectively. Decomposing these changes by workers' geographic mismatch, we see that in the short-run the biggest changes occur among geographically mismatched workers, which suggests that many of these marriages were sustained by the prospect of migration. However, in the long-run equilibrium, the change in the marriage market conditions induces an increase in this fraction for the geographically mismatched.

6 Conclusions

The interplay between marriage and mobility has the potential to affect a variety of economic outcomes. In this paper, we establish and quantify a series of important aspects of this interaction. First, we provide some evidence that individuals are forward-looking and take into account the labor-market costs due to reduced mobility when selecting into marriage.

	Divorce Rates				
	Baseline (%)	SR (%)	% Change	LR (%)	% Change
			All		
	0.54	0.69	+27.33	0.54	+0.40
Geographic Mismatch			Men		
1	0.44	0.53	+18.94	0.44	-1.69
2	0.71	0.95	+34.36	0.72	+2.30
3	0.87	1.19	+36.71	0.88	+1.88
4	0.57	0.79	+38.33	0.59	+2.82
5	1.40	2.29	+63.09	1.41	+0.41
Geographic Mismatch			Women		
1	0.46	0.55	+19.66	0.45	-1.42
2	0.74	1.02	+36.94	0.76	+2.37
3	0.88	1.14	+29.42	0.89	+0.80
4	0.57	0.87	+52.20	0.61	+5.98
5	1.28	2.16	+68.43	1.31	+1.63

Table 12: Divorce rates in the baseline and counterfactual scenario where migration is not allowed.

	Fraction of occupationally homogeneous marriages				
	Baseline (%)	SR (%)	% Change	LR (%)	% Change
			All		
	8.29	8.09	-2.37	8.05	-2.90
Geographic Mismatch			Men		
1	9.03	8.88	-1.63	8.73	-3.25
2	7.57	7.33	-3.17	7.36	-2.74
3	4.94	4.94	+0.09	5.15	+4.21
4	6.76	6.43	-4.87	6.87	+1.69
5	6.58	5.97	-9.25	7.17	+8.92
Geographic Mismatch			Women		
1	8.54	8.39	-1.81	8.19	-4.13
2	8.29	8.00	-3.55	8.04	-3.08
3	6.53	6.53	-0.07	6.96	+6.51
4	7.62	7.25	-4.95	7.90	+3.7
5	6.67	6.01	-9.87	7.48	+12.18

Table 13: Yearly fraction of new marriages involving spouses employed in the same occupation in the baseline and counterfactual scenario where migration is not allowed.

We exploit the geographic heterogeneity in wages across US cities for different occupations to build a measure of geographic mismatch meant to be a proxy for the value assigned by workers to future migration opportunities. We empirically show that geographically mis-

matched individuals (i.e. those workers who would gain the most from migrating) are less likely to enter marriage and, if married, more likely to divorce. Moreover, conditional on entering marriage, geographically mismatched individuals tend to marry similarly mismatched partners.

We rationalize these correlations through a stationary equilibrium, heterogeneous agents model with endogenous marriage, divorce, and migration. In the model, agents differ by their gender, their occupation, and their idiosyncratic preference for the city in which they live. Within a marriage, resources are allocated through Nash bargaining, where the outside option for both spouses is a costly divorce. The outside option of being single bears the full value of the migration option, while within marriage the migration value for one individual is reduced due to bargaining. The marital surplus for couples involving geographically mismatched spouses is then, on average, smaller and, thus, their marriage is less stable. Moreover, ceteris paribus, the marital surplus tends to be higher if the individual incentives for migration of the two spouses align.

The model is estimated through indirect inference to match some of the empirical relations established in the empirical analysis. The estimated model is used to study the effects of marriage on labor market outcomes and the geographic allocation of labor and, on the other side, the effect of migration on marriage market outcomes. We find that by entering an early marriage that restricts their mobility, individuals give up as much as 10% in wage growth over their working life. Moreover, we find that heterogeneity in local marriage market conditions constitutes an additional incentive for workers to relocate to more productive cities. We estimate that this effect increases the geographic concentration of labor in more productive cities, causing an increase in aggregate labor earning of 2.7%. Finally, we estimate that, without migration, the average yearly marriage probability would be 3.6% higher and the average yearly divorce rate 0.4% higher.

The analysis carried out in this paper opens up several interesting questions. On the one hand, it suggests that policies and technological changes that affect the value of marriage can, by the mechanism described in this paper, have implications for the geographic distribution of labor and, as a consequence, for aggregate productivity. On the other hand, all the economic forces that shape the geography of wages also affect the dynamics of family formation and, consequently, they have potential implications for all the other aspects that correlate with it, such as inequality (aggregate, local, and within-household) and fertility. We postpone the study of these questions to future research.

References

Aiyagari, S. Rao, Jeremy Greenwood, and Nezih Guner. 2000. "On the state of the union." *Journal of Political Economy*, 108(2): 213–244.

Alesina, Alberto, and Dani Rodrik. 1994. "Distributive politics and economic growth." *Quarterly Journal of Economics*, 109(2): 465–490.

Alonzo, Davide, and Giovanni Gallipoli. 2020. "The Distribution of Labor Market Surplus." *Working Paper.*

Attanasio, Orazio, Hamish Low, and Virginia Sánchez-Marcos. 2008. "Explaining changes in female labor supply in a life-cycle model." *American Economic Review*, 98(4): 1517–1552.

Autor, David H, and David Dorn. 2013. "The growth of low-skill service jobs and the polarization of the US Labor Market." *American Economic Review*, 103(5): 1553–1597.

Baum-Snow, Nathaniel, and Ronni Pavan. 2012. "Understanding the city size wage gap." *Review of Economic Studies*, 79(1): 88–127.

Beaudry, Paul, David A Green, and Benjamin Sand. 2012. "Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory." *Econometrica*, 80(3): 1063–1104.

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

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

Becker, Gary S. 1974*b*. "A Theory of Social Interactions." *Journal of Political Economy*, 82(6): 1063–1093.

Blanchard, Olivier J. 1985. "Debt, Deficits, and Finite Horizons." *Journal of Political Economy*, 93(2): 223–247.

Botticini, Maristella, and Aloysius Siow. 2007. "Are There Increasing Returns to Scale in Marriage Markets ?" *Unpublished Manuscript*, 1–34.

Braun, Christine, Charlie Nusbaum, and Peter Rupert. 2019. "Labor Market Dynamics and the Migration Behavior of Married Couples."

Cancian, Maria, and Deborah Reed. 1998. "Assessing the effects of wives' earnings on family income inequality." *Review of Economics and Statistics*, 80(1): 73–79.

Caucutt, Elizabeth M, Nezih Guner, and John Knowles. 2002. "Why Do Women Wait? Matching, Wage Inequality, and the Incentives for Fertility Delay." *Review of Economic Dynamics*, 5(4): 815–855.

Chiappori, Pierre-André. 1988. "Rational household labor supply." *Econometrica*, 56(1): 63–90.

Chiappori, Pierre-André. 1992. "Collective Labor Supply and Welfare." *Journal of Political Economy*, 100(3): 437–467.

Chiappori, Pierre-André, and Costas Meghir. 2015. "Intrahousehold Inequality." *Handbook of Income Distribution*, 2: 1369–1418.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss. 2017. "Partner Choice, Investment in Children, and the Marital College Premium." *American Economic Review*, 107(8): 2109–2167.

Chiappori, Pierre André, Monica Costa-Dias, and Costas Meghir. 2020. "Changes in Assortative Matching: Theory and Evidence for the US." *Cowles Fundation Discussion Paper No. 2226.*

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

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): 175–201.

Choo, Eugene, and Aloysius Siow. 2007. "Lifecycle marriage matching : Theory and evidence." *Working Paper.*

Compton, Janice, and Robert A. Pollak. 2007. "Why are power couples increasingly concentrated in large metropolitan areas?" *Journal of Labor Economics*, 25(3): 475–512.

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

Dahl, Gordon B. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica*, 70(6): 2367–2420.

Daly, Martin, Margo Wilson, and Vasdev Shawn. 2001. "Income Inequality and Homicide Rates in Canada and the United States." *Canadian Journal of Criminology*, 43(2): 219–236.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum. 2019. "Matching in Cities."

D'Costa, Sabine, and Henry G Overman. 2014. "The urban wage growth premium: Sorting or learning?" *Regional Science and Urban Economics*, 48: 168–179.

De la Roca, Jorge, and Diego Puga. 2017. "Learning byworking in big cities." *Review of Economic Studies*, 84(1): 106–142.

Deming, David, and Lisa B Kahn. 2018. "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals." *Journal of Labor Economics*, 36(S1): S337–S369.

Devereux, Kevin, and Laura Turner. 2016. "Divorce, Remarriage, and Fertility with On-the-Marriage Search." *Working Paper.*

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

Diaz-Gimenez, Javier, and Eugenio Giolito. 2013. "Accounting for the Timing of First Marriage." *International Economic Review*, 54(1): 135–158.

Dorn, David. 2009. "Essays on Inequality, Spatial Interaction, and the Demand for Skills." PhD diss. University of St. Gallen.

Edlund, Lena. 2005. "Sex and the city." *Scandinavian Journal of Economics*, 107(1): 25–44.

Eika, Lasse, Magne Mogstad, and Basit Zafar. 2019. "Educational Assortative Mating and Household Income Inequality." *Journal of Political Economy*, 127(6): 2795–2835.

Fajnzylber, Pablo, Daniel Lederman, and Norman Loayza. 2002. "Inequality and violent crime." *Journal of Law and Economics*, 45(1): 1–39.

Fernandez, Raquel, and Richard Rogerson. 2001. "Sorting and Long-Run Inequality." *Quarterly Journal of Economics*, 116(November): 1305–1341.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2020. "Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]." *Minneapolis, MN: IPUMS, 2018.*

Frank, Robert H. 1978. "Why women earn less: The theory and estimation of differential over qualification." *American Economic Review*, 68(3): 360–373.

Galichon, Alfred, Scott Duke Kominers, and Simon Weber. 2019. "An Empirical Framework for Matching with Imperfectly Transferable Utility." *Journal of Political Economy*, 127(6): 2875–2925.

Gallipoli, Giovanni, and Laura Turner. 2011. "Social Security, Endogenous Retirement and Intrahousehold Cooperation." *Working Paper.*

Gathmann, Christina, and Uta Schönberg. 2010. "How General Is Human Capital? A Task-Based Approach." *Journal of Labor Economics*, 28(1): 1–49.

Gautier, Pieter A., Michael Svarer, and Coen N. Teulings. 2010. "Marriage and the city: Search frictions and sorting of singles." *Journal of Urban Economics*, 67(2): 206–218.

Gayle, George-Levi, and Andrew Shephard. 2019. "Optimal Taxation, Marriage, Home Production, and Family Labor Supply." *Econometrica*, 87(1): 291–326.

Gemici, Ahu. 2016. "Family migration and labor market outcomes." *Manuscript.*

Gervais, Martin, Nir Jaimovich, Henry E. Siu, and Yaniv Yedid-Levi. 2016. "What Should I Be When I Grow Up? Occupations and Unemployment over the Life Cycle." *Journal of Monetary Economics*, 83: 54–70.

Glaeser, Edward L., Matt Resseger, and Kristina Tobio. 2009. "Inequality in cities." *Journal of Regional Science*, 49(4): 617–646.

Gourieroux, C., A. Monfort, and E. Renault. 1993. "Indirect Inference." *Journal of Applied Econometrics*, 8: S85–S118.

Goussé, Marion, Nicolas Jacquemet, and Jean-Marc Robin. 2017. "Marriage, Labor Supply, and Home Production." *Econometrica*, 85(6): 1873–1919.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. 2005. "Engines of liberation." *Review of Economic Studies*, 72(1): 109–133.

Greenwood, Jeremy, Nezih Guner, and John A. Knowles. 2000. "Women on welfare: A macroeconomic analysis." *American Economic Review*, 90(2): 383–388.

Greenwood, Jeremy, Nezih Guner, and John A. Knowles. 2003. "More on marriage, fertility, and the distribution of income." *International Economic Review*, 44(3): 827–862.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2014. "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review: Papers & Proceedings*, 104(5): 348–353.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2016. "Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation." *American Economic Journal: Macroeconomics*, 8(1): 1–41.

Guler, Bulent, Fatih Guvenen, and G.L. Violante. 2011. "Joint-search theory: New opportunities and new frictions." *Journal of Monetary Economics*, 59(4): 352–369.

Ibes, Dorothy C. 2015. "A multi-dimensional classification and equity analysis of an urban park system: A novel methodology and case study application." *Landscape and Urban Planning*, 137: 122–137.

Kaplan, Greg, and Sam Schulhofer-Wohl. 2017. "Understanding the Long-Run Decline in Interstate Migration." *International Economic Review*, 58(1): 57–94.

Kennan, John, and James R Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79(1): 211–251.

Knowles, John A. 2013. "Why are married men working so much? An aggregate analysis of intra-household bargaining and labour supply." *Review of Economic Studies*, 80(3): 1055–1085.

Lazear, Edward P. 2009. "Firm-Specific Human Capital: A Skill-Weights Approach." *Journal of Political Economy*, 117(5): 914–940.

Low, Hamish, Costas Meghir, Luigi Pistaferri, and Alessandra Voena. 2018. "Marriage, Labor Supply and the Dynamics of the Social Safety Net." *NBER working paper*.

Lundberg, Shelly, and Ra Pollak. 1996. "Bargaining and distribution in marriage." *The Journal of Economic Perspectives*, 10(4): 139–158.

Lundberg, Shelly, and Robert Pollak. 2013. "Cohabitation and the Uneven Retreat from Marriage in the U.S., 1950-2010."

Mazzocco, Maurizio, Claudia Ruiz, and Shintaro Yamaguchi. 2013. "Labor Supply , Wealth Dynamics, and Marriage Decisions." *Working Paper*, 1–51.

McFadden, Daniel. 1974. "Conditional logit analysis of qualitative choice behavior." In *Frontiers in Econometrics*. 105–142.

Melo, Patricia C, Daniel J Graham, and Robert B Noland. 2009. "A meta-analysis of estimates of urban agglomeration economies." *Regional Science and Urban Economics*, 39(3): 332–342.

Menzio, Guido, Irina A Telyukova, and Ludo Visschers. 2016. "Directed search over the life cycle." *Review of Economic Dynamics*, 19: 38–62.

Mincer, Jacob. 1978. "Family Migration Decisions." *Journal of Political Economy*, 86(5): 749–773.

Moretti, Enrico. 2004*a*. "Estimating the social return to higher education : evidence from longitudinal and repeated cross-sectional data." *Journal of Econometrics*, 121(1): 175–212.

Moretti, Enrico. 2004*b*. "Human capital externalities in cities." *Handbook of regional and urban economics*, 4: 2243–2291.

Moretti, Enrico. 2004*c*. "Workers' Education , Spillovers , and Productivity : Evidence from Plant-Level Production Functions." *American Economic Review*, 94(3): 656–690.

Moretti, Enrico. 2012. *The New Geography of Jobs*.

Moretti, Enrico. 2013. "Real wage inequality." *American Economic Journal: Applied Economics*, 5(1): 65–103.

Papageorgiou, Thoeodore. 2013. "Learning Your Comparative Advantages." *Review of Economic Studies*, 81(3): 1263–1295.

Persson, Torsten, and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *American Economic Review*, 84(3): 600–621.

Poletaev, Maxim, and Chris Robinson. 2008. "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000." *Journal of Labor Economics*, 26(3): 387–420.

Rouwenhorst, Geert. 1995. "Asset pricing implications of equilibrium business cycle models." In *Frontiers of business cycle research.* Chapter 10. Princeton University Press.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. "IPUMS USA: Version 10.0 [dataset]." *Minneapolis, MN: IPUMS.*

Schwartz, Christine R. 2010. "Earnings inequality and the changing association between spouses' earnings." *American Journal of Sociology*, 115(5): 1524–1557.

Shephard, Andrew. 2019. "Marriage market dynamics, gender, and the age gap." *Working Paper*, 1–37.

Siow, Aloysius. 2015. "Assortative Mating in the Marriage Market." *International Encyclopedia of the Social & Behavioral Sciences: Second Edition.*

Swensen, Isaac, Jason M. Lindo, and Krishna Regmi. 2020. "Stable Income, Stable Family." *NBER Working Paper.*

Voena, Alessandra. 2015. "Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples?" *American Economic Review*, 105(8): 2295–2332.

Yamaguchi, Shintaro. 2012. "Tasks and Heterogeneous Human Capital." *Journal of Labor Economics*, 30(1): 1–53.

Yeung, Ka Yee, and Walter L. Ruzzo. 2001. "An Empirical study on Principal Component Analysis for Clustering Gene Expression Data." *Bioinformatics*, 17: 763–774.

A Cohabitation

In the main text, we argue that cohabiting couples might behave differently in comparison to legally married couples because of the less stringent nature of cohabitation versus marriage. In this appendix, we present evidence that this is indeed the case, showing that cohabiting couples do not display different migration rates than single households. The ACS sample allows us to distinguish between simple housemates from cohabiting, non-married couples thanks to a variable ("related") that defines the relationship of household members with the head. One of the possible values for this variable is "unmarried partner." Among all non-married households, about 13.9% report the head being in an unmarried relationship. We use this information to compare the migration rates of singles to the subset of unmarried couples. The comparison is carried out estimating a logit model for the probability of divorce of the following form:

$$Pr\left(Y_{i,t}=1\right)=\varphi\left(\beta_0+\beta_1 Cohabiting_{i,t}+\boldsymbol{\beta}_2\boldsymbol{X}_{i,t}\right)$$

where $Cohabiting_{i,t}$ is a dummy that indicates whether individual i has a partner cohabiting with them and $\boldsymbol{X}_{i,t}$ contains controls for age, education, and the presence of children. Table 14 reports the estimated coefficient (in odds-ratio form) on the $Cohabiting_{i,t}$ dummy. It clearly shows that there is no statistically significant difference between the migration rates of singles and unmarried couples.

	Yearly Probability of Migration (Odds-ratios)			
	Men		Women	
Cohabiting	1.178***	1.036	1.546***	1.082
	(0.0641)	(0.0570)	(0.0828)	(0.0592)
Observations	278,926	278,926	240,392	240,392
Demographic Controls	NO	YES	NO	YES

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14: Regression-based odds ratios for the probability of migration of cohabiting partners vs. singles. Demographic controls include age, education, and presence of children.

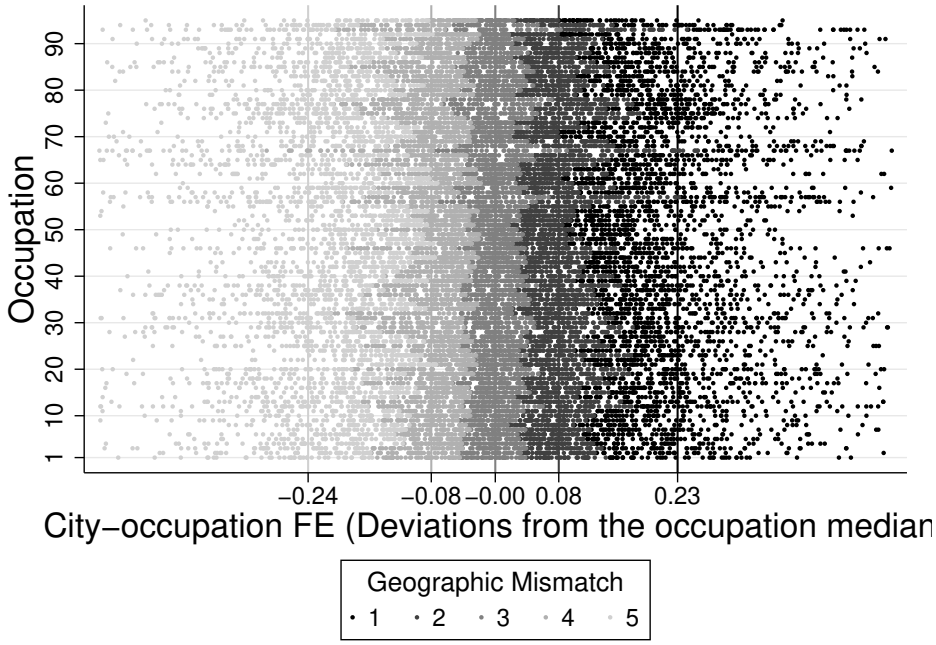

Figure 16: Estimated city-occupation fixed effects (deviations from the occupation median). Different shades of gray represent the five levels of geographic mismatch. The vertical lines report the averages for each group.

B Tables and Additional Figures

In this section, we report all the extra figures, tables, and robustness checks relative to the main findings of section 2. Additional robustness checks are reported in the following sections.

City-occupation Fixed Effects and Rankings. The data show substantial geographic heterogeneity in occupational returns. The standard deviation in the estimated city-occupation fixed effect, which captures the dispersion in log-hourly wages across cities, is 16% across occupations, ranging from a minimum of 9% to a maximum of 35%. Figure 16 shows the estimated city-occupation fixed effects as deviations from the occupation median. The shades of gray indicate the corresponding level of geographic mismatch assigned by our procedure and the vertical lines correspond to the mean values of this deviation from the median for the 5-levels of geographic mismatch.

The estimated fixed effects are used to construct occupation-specific rankings. Figure 17 is a graphical representation of these rankings. On the y-axis, we have cities ordered by their average ranking across occupations. The San Francisco-Oakland-Hayward MSA, for instance, is the one that has the highest average ranking across occupations. Each point in

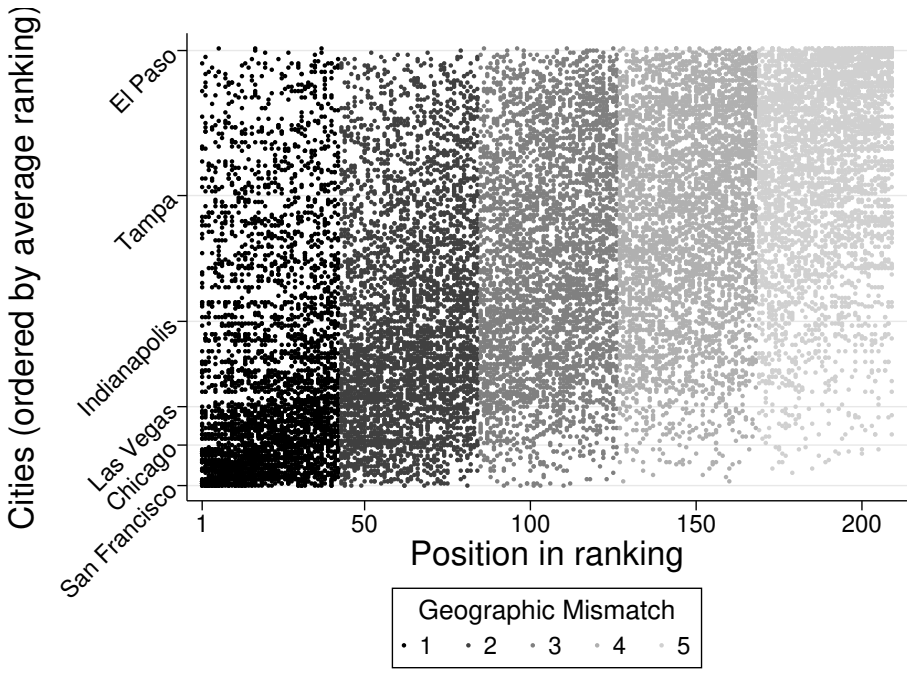

Figure 17: Graphic representation of occupation-specific city rankings. On the y-axis, cities are ordered based on average ranking across occupations. Each point in the graph corresponds to a particular city-occupation pair and it is placed in correspondence to its ranking on the x-axis.

the graph corresponds to a particular city-occupation pair and it is placed in correspondence to its ranking on the x-axis. The color coding indicates the level of geographic mismatch associated with each city-occupation pair. The graph shows that, while it is true that some cities tend to rank higher for many occupations, there is a lot of variation that ensures that there are geographically mismatched individuals in each city.

Migration Tables 15 and 16 report the estimates for the migration equation for men and women respectively. The second column in both tables corresponds to the baseline model represented in figure 1. The first column corresponds to the migration regression used as the auxiliary model in the estimation of the model through indirect inference. The only difference with the former comes from the lack of the education and the dummy for the presence of children (both are sources of heterogeneity that are not modeled) and the absence of year dummies. In the remaining columns, we perform robustness checks including occupation fixed effects and state (of origin) fixed effect, separately and together. Moreover, in the last two columns, we estimate the model splitting the sample by education: same college and above (labeled as "College") and high-school graduates and below ("Non-College"). The results are clearly robust across columns. Interestingly, the inclusion of the state fixed

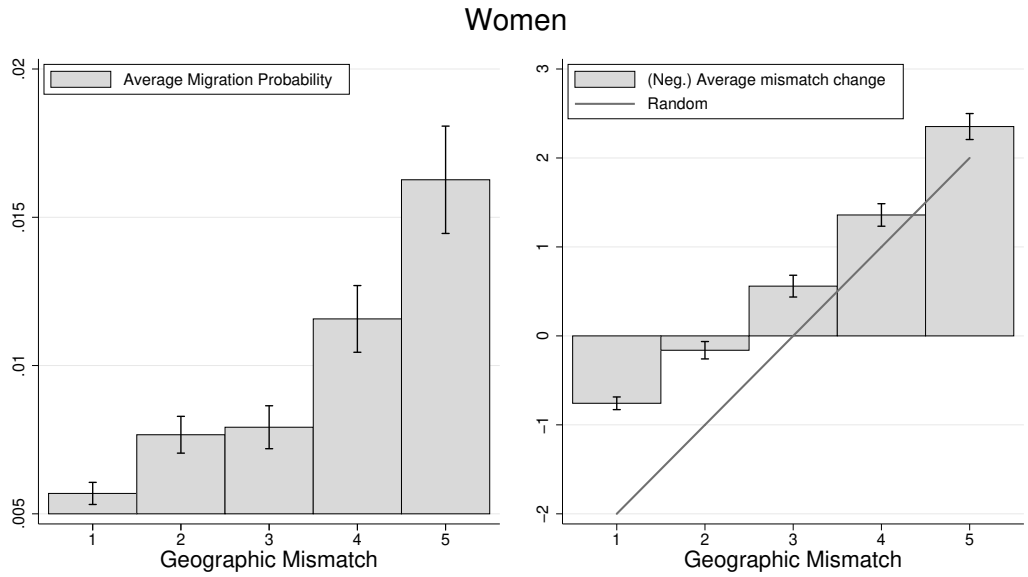

Figure 18: Migration rates and average (negative) change in geographic mismatch conditional on migration happening for single women households.

effects increases the steepness of the relation between migration probabilities and workers' mismatch levels. The data does not provide enough power to allow us to control for city-specific fixed effects separately in the regression for men and women. Nevertheless, given the similarity in the results for the two groups, we can estimate a pooled regression in which we can include city fixed effects (we also include a gender dummy). As shown by table 17, even after controlling for city fixed effects, the qualitative relationship between the probability of migration and geographic mismatch persists.

Figure 18 is the women counterpart of figure 1 and shows the migration rates (corresponding to the estimates in the second column of Table 16) and the average negative change in geographic mismatch conditional on migration.

Marriage and Divorce Tables 18 and 19 show the estimates of the marriage equation on the sample of men and women respectively. All the columns follow the same logic as in the migration table. We see that the results are robust to the inclusion of occupation and state fixed effects separately for men. Also, the results are robust to spitting the sample by education. For women, in all the specifications many coefficients do not achieve statistical significance just like in the baseline estimate. Only the coefficient on the dummy for mismatch level 5 is often significant, but point estimates are in line with our expectations. Even if not directly related to the topic of this paper, an interesting fact emerges from the table. The inclusion of occupation fixed effects reverses the effect of education on marriage

	Yearly Probability of Migration (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.315***	1.370***	1.397***	1.464***	1.498***	1.053
	(0.0635)	(0.0662)	(0.0690)	(0.0824)	(0.0849)	(0.0955)
Mismatch 3	1.417***	1.502***	1.523***	1.836***	1.661***	1.140
	(0.0720)	(0.0766)	(0.0807)	(0.116)	(0.101)	(0.105)
Mismatch 4	1.542***	1.687***	1.765***	2.033***	1.933***	1.218**
	(0.0868)	(0.0955)	(0.105)	(0.143)	(0.133)	(0.118)
Mismatch 5	2.068***	2.339***	2.481***	3.026***	3.095***	1.392***
	(0.126)	(0.145)	(0.162)	(0.239)	(0.230)	(0.146)
Wage	1.124***	1.009	0.930**	1.002	1.051	0.953
	(0.0335)	(0.0303)	(0.0286)	(0.0312)	(0.0376)	(0.0544)
Wage dispersion	6.412***	3.916***		3.427***	6.165***	2.113
	(2.582)	(1.631)		(1.445)	(3.069)	(1.596)
Children		0.761***	0.804***	0.769***	0.658***	0.866
		(0.0512)	(0.0545)	(0.0523)	(0.0708)	(0.0761)
College dummy		1.752***	1.440***	1.697***		
		(0.0697)	(0.0643)	(0.0677)		
Observations	453,245	453,245	453,245	437,069	262,570	190,675
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 15: Estimates of the migration equation for men. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

for men. Within occupations, educated men are less likely to marry. This clashes with the common view that well-educated men are more likely to marry because they are more "competitive" in the marriage market than less-educated men. The opposite is true for women. The inclusion of occupation fixed effects seems to increase the correlation of education with marriage probabilities. This suggests that highly educated women are more likely to work in occupations that make them less likely to marry. Further investigations are postponed to future research.

Figure 19 is the women counterpart of Figure 4.

Figure 20 shows the values of our baseline measure of assortative mating (eq. 4) consid-

	Yearly Probability of Migration (Odds-ratios) - Women					
	All Women				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.293***	1.352***	1.333***	1.718***	1.412***	1.066
	(0.0701)	(0.0732)	(0.0741)	(0.110)	(0.0829)	(0.147)
Mismatch 3	1.335***	1.398***	1.360***	2.050***	1.476***	1.052
	(0.0783)	(0.0819)	(0.0826)	(0.149)	(0.0937)	(0.158)
Mismatch 4	1.915***	2.054***	1.977***	2.869***	2.193***	1.511***
	(0.119)	(0.127)	(0.130)	(0.222)	(0.148)	(0.225)
Mismatch 5	2.707***	2.909***	2.852***	4.179***	2.974***	2.525***
	(0.187)	(0.200)	(0.206)	(0.368)	(0.230)	(0.389)
Wage	1.150***	0.963	0.905**	0.948	0.957	1.021
	(0.0396)	(0.0349)	(0.0352)	(0.0354)	(0.0387)	(0.0873)
Wage dispersion	19.32***	9.330***		8.941***	13.87***	1.465
	(9.348)	(4.671)		(4.560)	(7.606)	(1.809)
Children		0.593***	0.617***	0.592***	0.523***	0.730***
		(0.0359)	(0.0378)	(0.0362)	(0.0418)	(0.0749)
College dummy		1.932***	1.777***	1.877***		
		(0.107)	(0.107)	(0.106)		
Observations	352,015	352,015	350,484	340,027	249,217	102,798
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 16: Estimates of the migration equation for women. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

ering all married couples instead of only newlyweds. For comparison, the dotted line is the baseline measure from figure 5. The graph shows roughly the same patterns as the baseline. The increase in the degree of marital sorting along the diagonal is lower than in the baseline and the level itself is always below. This is consistent with and can be interpreted as an effect of the migration patterns described in the paper.

Figures 21 and 22 show the results from computing a conditional version of the assortative mating measure in eq. (4). In this version, we compare the observed distribution of marriages across two dimensions, geographic mismatch and the characteristics we wish to control for, to the one we would obtain if marriage was random with respect to geographic mismatch but

	Yearly Probability of Migration (Odds-ratios) - All All Men and Women				
Mismatch 1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.307***	1.364***	1.369***	1.556***	1.070
	(0.0471)	(0.0491)	(0.0505)	(0.0657)	(0.0483)
Mismatch 3	1.383***	1.460***	1.450***	1.915***	1.163***
	(0.0530)	(0.0561)	(0.0578)	(0.0912)	(0.0591)
Mismatch 4	1.693***	1.840***	1.852***	2.334***	1.172***
	(0.0706)	(0.0768)	(0.0818)	(0.122)	(0.0650)
Mismatch 5	2.309***	2.571***	2.632***	3.438***	1.248***
	(0.106)	(0.118)	(0.128)	(0.202)	(0.0770)
Wage	1.134***	0.992	0.922***	0.981	1.001
	(0.0256)	(0.0229)	(0.0222)	(0.0234)	(0.0245)
Wage dispersion	9.920***	5.666***		5.263***	9.569***
	(3.084)	(1.813)		(1.715)	(3.111)
Children		0.656***	0.686***	0.656***	0.626***
		(0.0293)	(0.0314)	(0.0300)	(0.0288)
College dummy		1.818***	1.561***	1.759***	1.752***
		(0.0576)	(0.0553)	(0.0568)	(0.0564)
Observations	805,260	805,260	805,260	777,096	575,973
Year dummies	NO	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO
State FE	NO	NO	NO	YES	NO
MSA FE	NO	NO	NO	NO	YES

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 17: Estimates of the pooled migration equation. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

not with respect to the controlling characteristics. Let H and W denote the characteristic we wish to control for husbands and wives respectively. The new measure is given by

$$s_{h,w} = \frac{P(h, w, H, W)}{\sum_c P(h|H, c)P(w|W, c)P(H, W|c)P(c)} \tag{28}$$

Figure 21, reports the computed statistics controlling for education. In particular, it shows marital sorting conditional on both spouses having the same education. In figure 22, we condition on both spouses belonging to the same age group. Due to data limitations, both are computed on the full set of married couples and not on the subset of newlyweds as we do

	Yearly Probability of Marriage (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	0.937***	0.920***	0.947***	0.933***	0.929***	0.886***
	(0.0169)	(0.0168)	(0.0180)	(0.0188)	(0.0194)	(0.0338)
Mismatch 3	0.911***	0.891***	0.907***	0.904***	0.912***	0.837***
	(0.0185)	(0.0183)	(0.0196)	(0.0210)	(0.0218)	(0.0342)
Mismatch 4	0.893***	0.868***	0.931***	0.885***	0.860***	0.882***
	(0.0218)	(0.0214)	(0.0247)	(0.0242)	(0.0263)	(0.0380)
Mismatch 5	0.801***	0.766***	0.857***	0.781***	0.837***	0.681***
	(0.0253)	(0.0247)	(0.0296)	(0.0273)	(0.0345)	(0.0362)
City size	0.989*	0.996	0.990	1.000	0.999	1.000
	(0.00640)	(0.00650)	(0.00688)	(0.00760)	(0.00784)	(0.0122)
Sex ratio	0.0922***	0.0811***	0.154***	0.563	0.102***	0.0612***
	(0.0216)	(0.0192)	(0.0364)	(0.212)	(0.0280)	(0.0291)
Wage	1.264***	1.246***	1.190***	1.246***	1.226***	1.351***
	(0.0135)	(0.0139)	(0.0146)	(0.0140)	(0.0157)	(0.0313)
Wage dispersion	1.840***	1.881***		1.902***	1.870***	2.048**
	(0.320)	(0.332)		(0.336)	(0.388)	(0.687)
Children		2.838***	3.309***	2.866***	3.061***	2.644***
		(0.0520)	(0.0634)	(0.0527)	(0.0724)	(0.0756)
College dummy		1.564***	0.954**	1.561***		
		(0.0278)	(0.0195)	(0.0278)		
Observations	308,457	308,457	308,457	308,457	214,564	93,893
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18: Estimates of the marriage equation for men. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

in the baseline. The positive relationship between the level of geographic mismatch and the degree of assortative mating is robust to controlling for education or age. It is worth noting that, when controlling for age, younger couples, who are more likely to be newlyweds and have higher migration rates, display steeper profiles than older couples. This is consistent with the mechanism proposed in the paper.

Tables 20 and 21 show the results for the probability of marrying within an occupation.

	Yearly Probability of Marriage (Odds-ratios) - Women					
	All Women				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	0.987	0.992	1.009	1.002	0.976	1.048
	(0.0181)	(0.0182)	(0.0193)	(0.0209)	(0.0196)	(0.0485)
Mismatch 3	1.000	1.001	0.990	1.007	0.987	1.055
	(0.0204)	(0.0205)	(0.0214)	(0.0243)	(0.0222)	(0.0531)
Mismatch 4	0.977	0.979	0.945**	0.980	0.966	1.019
	(0.0248)	(0.0250)	(0.0257)	(0.0283)	(0.0275)	(0.0594)
Mismatch 5	0.937*	0.939*	0.866***	0.947	0.930*	0.979
	(0.0321)	(0.0323)	(0.0314)	(0.0357)	(0.0362)	(0.0728)
City size	0.988*	0.989	0.990	0.989	0.980***	1.029*
	(0.00680)	(0.00686)	(0.00700)	(0.00794)	(0.00756)	(0.0164)
Sex ratio	0.768	0.845	0.707	5.670***	0.673	2.965**
	(0.170)	(0.187)	(0.163)	(2.104)	(0.164)	(1.558)
Wage	1.313***	1.237***	1.218***	1.240***	1.255***	1.227***
	(0.0154)	(0.0155)	(0.0169)	(0.0156)	(0.0175)	(0.0360)
Wage dispersion	0.551***	0.503***		0.540***	0.401***	3.109**
	(0.104)	(0.0962)		(0.104)	(0.0832)	(1.509)
Children		1.143***	1.122***	1.149***	1.258***	0.932*
		(0.0211)	(0.0217)	(0.0213)	(0.0261)	(0.0334)
College dummy		1.466***	1.703***	1.464***		
		(0.0313)	(0.0404)	(0.0314)		
Observations	275,612	275,612	275,612	275,612	214,529	61,083
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 19: Estimates of the marriage equation for women. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

Columns are as for the other estimates, where the second columns are the ones depicted in figure 6. All the results are robust to the inclusion of occupation or state fixed effects (even though not all the estimates remain significant in all the specifications). The results still hold when estimating the probability model on the subset of college graduates while, due to the small sample size, significance is lost in almost all the coefficients on the dummies for geographic mismatch when estimating on the lower education group.

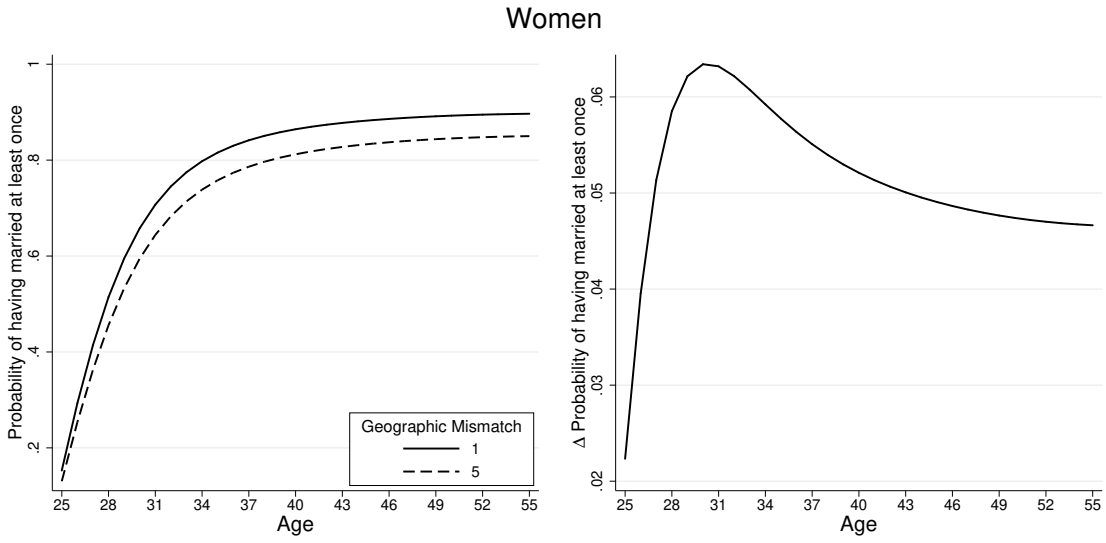

Figure 19: Estimated probability of having married at least once as a function of age for women with the lowest and highest levels of geographic mismatch. The left panel shows the estimated probabilities while the right panel the difference between the two.

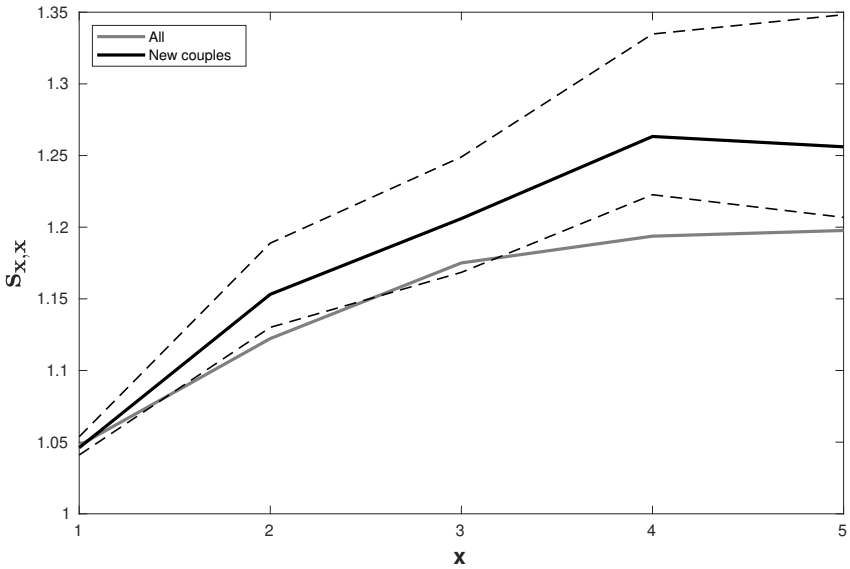

Figure 20: Our measure of marital sorting as a function of the spouses' geographic mismatch for all couples. The figure focuses on the diagonal values and includes 95% confidence intervals and the baseline measure (dashed line).

Finally, tables 22 and 23 report the results from the estimation of the divorce equation for men and women respectively. Columns are as described previously. Also in this case, we see that the results are robust.

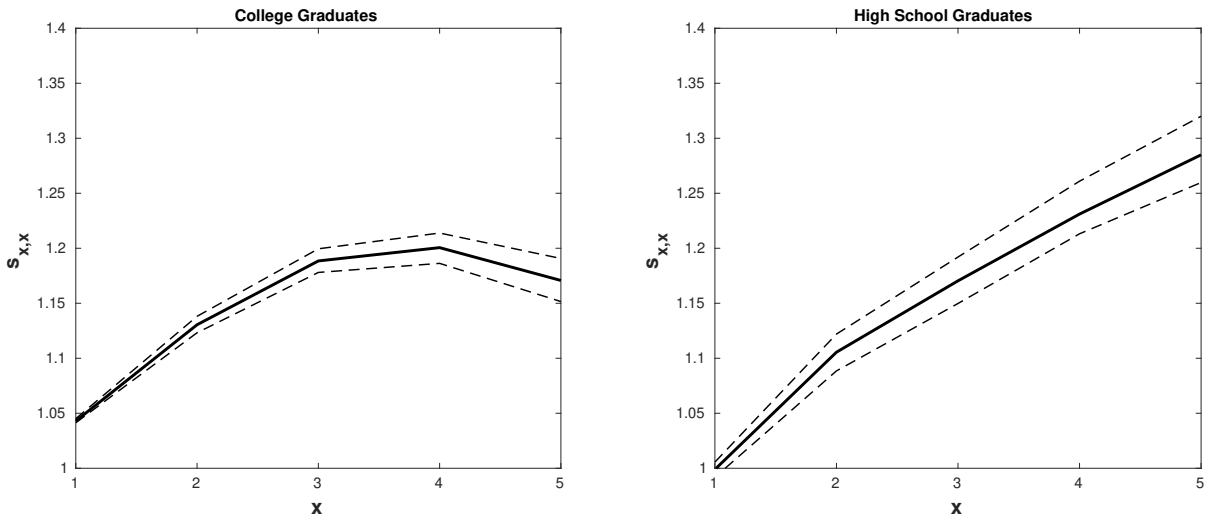

Figure 21: Our measure of marital sorting as a function of the spouses' geographic mismatch for all couples, conditional on both spouses having the same education.

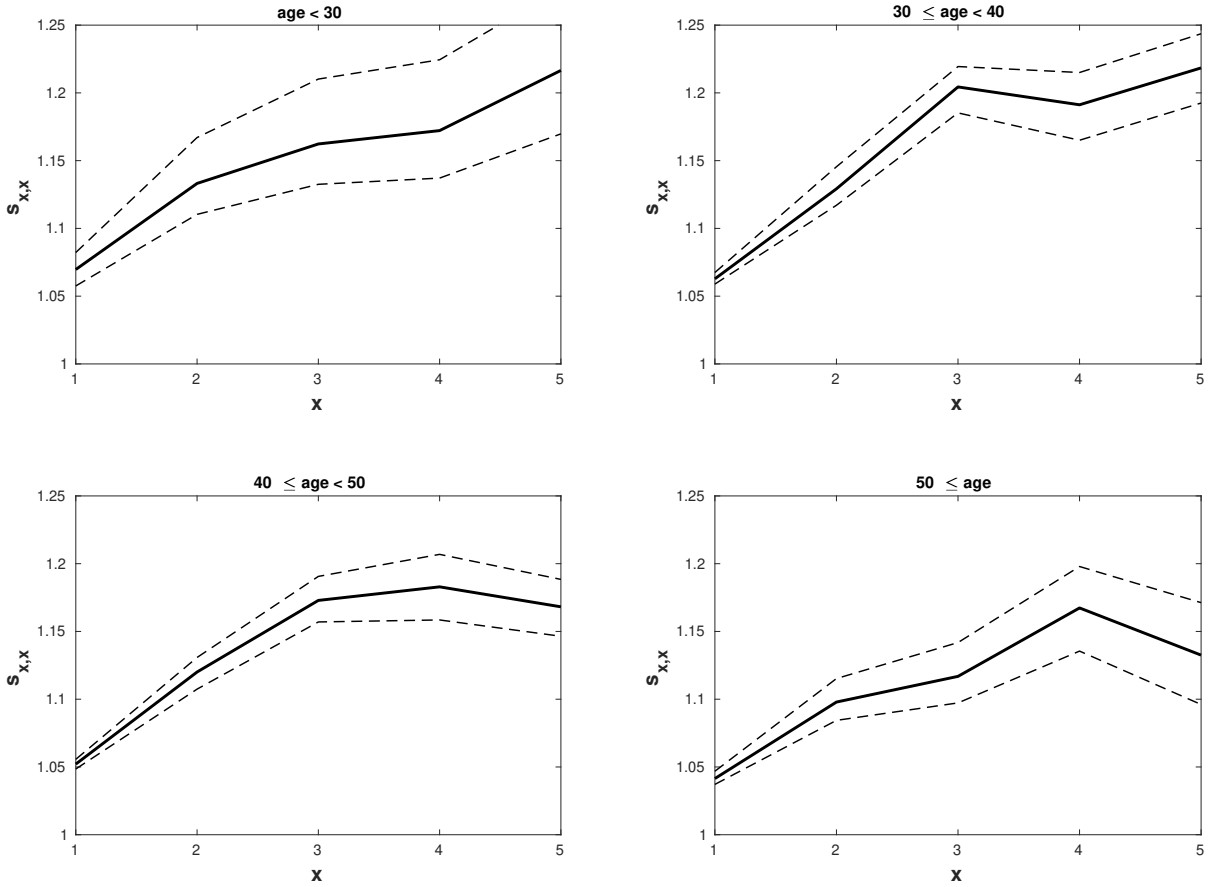

Figure 22: Our measure of marital sorting as a function of the spouses' geographic mismatch for all couples, conditional on both spouses belonging to the same age group.

	Probability of Marriage Within Occupation (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.041	1.058	1.043	1.083	1.041	1.148
	(0.0653)	(0.0665)	(0.0671)	(0.0784)	(0.0703)	(0.201)
Mismatch 3	1.259***	1.287***	1.221***	1.320***	1.296***	1.209
	(0.0852)	(0.0873)	(0.0850)	(0.105)	(0.0950)	(0.222)
Mismatch 4	1.284***	1.348***	1.176*	1.374***	1.447***	0.995
	(0.108)	(0.114)	(0.104)	(0.132)	(0.135)	(0.208)
Mismatch 5	1.445***	1.522***	1.221*	1.543***	1.670***	1.103
	(0.156)	(0.165)	(0.139)	(0.185)	(0.206)	(0.262)
Other sex fraction	1.437***	1.436***	1.920***	1.431***	1.410***	1.612***
	(0.0334)	(0.0336)	(0.0782)	(0.0336)	(0.0361)	(0.0958)
College		0.811	0.850	0.814		
		(0.121)	(0.135)	(0.122)		
College, spouse		0.506***	0.505***	0.510***	1.452***	0.522***
		(0.0654)	(0.0676)	(0.0659)	(0.182)	(0.0709)
College interaction		2.914***	2.465***	2.920***		
		(0.512)	(0.450)	(0.514)		
Wage	0.352***	0.328***	0.370***	0.331***	0.338***	0.357**
	(0.0816)	(0.0739)	(0.0802)	(0.0746)	(0.102)	(0.176)
Wage, spouse	0.440***	0.418***	0.472***	0.424***	0.455**	0.325**
	(0.114)	(0.106)	(0.117)	(0.108)	(0.155)	(0.168)
Wage interaction	1.446***	1.454***	1.391***	1.446***	1.433***	1.468**
	(0.111)	(0.108)	(0.101)	(0.108)	(0.141)	(0.257)
Observations	21,885	21,885	21,643	21,868	17,389	4,496
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 20: Estimates of the equation for marrying within occupation for men. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

ACS and CPS occupational structure In figure 24, we compare the gender-specific occupational distributions obtained from the ACS sample (used for the empirical analysis) to the corresponding stationary distributions obtained from the transition matrices computed from the CPS (fed to the model). In the figure, each point represents one of the 95

	Probability of Marriage Within Occupation (Odds-ratios) - Women					
	All Women				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.025	1.040	1.030	1.033	1.070	0.779
	(0.0638)	(0.0651)	(0.0660)	(0.0765)	(0.0700)	(0.162)
Mismatch 3	1.156**	1.185**	1.062	1.157*	1.184**	1.123
	(0.0770)	(0.0792)	(0.0723)	(0.0955)	(0.0837)	(0.234)
Mismatch 4	1.293***	1.321***	1.178*	1.321***	1.328***	1.191
	(0.108)	(0.111)	(0.103)	(0.135)	(0.120)	(0.283)
Mismatch 5	1.593***	1.647***	1.358***	1.612***	1.583***	1.632*
	(0.174)	(0.180)	(0.153)	(0.199)	(0.195)	(0.420)
Other sex fraction	1.447***	1.443***	1.461***	1.435***	1.399***	1.907***
	(0.0332)	(0.0332)	(0.0343)	(0.0331)	(0.0334)	(0.163)
College		0.557***	0.476***	0.565***		
		(0.0714)	(0.0633)	(0.0725)		
College, spouse		0.911	0.849	0.923	2.408***	0.949
		(0.135)	(0.131)	(0.137)	(0.225)	(0.148)
College interaction		2.679***	2.355***	2.652***		
		(0.467)	(0.420)	(0.463)		
Wage	0.532**	0.481***	0.478***	0.483***	0.532**	0.361***
	(0.131)	(0.118)	(0.114)	(0.118)	(0.152)	(0.140)
Wage, spouse	0.390***	0.350***	0.372***	0.352***	0.397***	0.218***
	(0.0855)	(0.0758)	(0.0773)	(0.0763)	(0.101)	(0.0812)
Wage interaction	1.377***	1.397***	1.374***	1.393***	1.353***	1.578***
	(0.0998)	(0.1000)	(0.0952)	(0.0998)	(0.111)	(0.220)
Observations	22,422	22,422	21,760	22,404	19,395	3,027
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 21: Estimates of the equation for marrying within occupation for women. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

occupations. A point lying on the 45° line implies identical shares in the two samples for the corresponding occupation. Overall, for both men and women, the two distributions are quite similar. In both cases, the major discrepancy is given by occupation 178, "Lawyers and judges" (see appendix J) which is overestimated in the CPS compared to the ACS sample.

	Yearly Probability of Divorce (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.101***	1.103***	1.072***	1.066***	1.114***	1.075**
	(0.0190)	(0.0193)	(0.0190)	(0.0208)	(0.0248)	(0.0309)
Mismatch 3	1.143***	1.139***	1.088***	1.076***	1.160***	1.095***
	(0.0208)	(0.0210)	(0.0204)	(0.0227)	(0.0278)	(0.0321)
Mismatch 4	1.192***	1.188***	1.086***	1.116***	1.220***	1.125***
	(0.0242)	(0.0245)	(0.0229)	(0.0262)	(0.0343)	(0.0349)
Mismatch 5	1.170***	1.181***	1.067**	1.103***	1.199***	1.122***
	(0.0285)	(0.0292)	(0.0269)	(0.0305)	(0.0429)	(0.0394)
Wage	0.745***	0.833***	0.858***	0.834***	0.866***	0.767***
	(0.00712)	(0.00902)	(0.00968)	(0.00901)	(0.0118)	(0.0133)
Wage dispersion	0.352***	0.450***		0.464***	0.441***	0.511***
	(0.0608)	(0.0774)		(0.0802)	(0.106)	(0.127)
College dummy		0.697***	0.921***	0.696***		
		(0.00992)	(0.0140)	(0.00993)		
Children		0.173***	0.173***	0.173***	0.170***	0.160***
		(0.00248)	(0.00250)	(0.00249)	(0.00322)	(0.00349)
Observations	947,971	947,971	947,971	947,971	636,433	311,538
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 22: Estimates of the divorce equation for men. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

Idetification and Estimation Figure 25 graphically illustrates how identification of the parameters is achieved. Each panel plots the value of one identifying variable as a function of different values of the parameter. For each parameter, we consider percentage deviations from the estimated value. In each panel, the black line refers to the parameter identified by the variable in question (as described in the body of this paper), while gray lines correspond to the other variables.

Figure 26 and table 24 are the analogous of figure 9 and table 8 and compare simulated and empirical equivalents of the target moments computed for women.

Figure 27 plots the same data as figure 10 but on an absolute scale.

	Yearly Probability of Divorce (Odds-ratios) - Women					
	All Women				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.073***	1.056***	1.025	1.014	1.104***	0.958
	(0.0173)	(0.0171)	(0.0167)	(0.0188)	(0.0218)	(0.0271)
Mismatch 3	1.076***	1.068***	1.054***	0.994	1.097***	0.999
	(0.0184)	(0.0184)	(0.0183)	(0.0201)	(0.0231)	(0.0298)
Mismatch 4	1.104***	1.091***	1.080***	1.023	1.116***	1.029
	(0.0216)	(0.0214)	(0.0214)	(0.0234)	(0.0275)	(0.0336)
Mismatch 5	1.085***	1.075***	1.105***	0.994	1.114***	1.000
	(0.0258)	(0.0257)	(0.0266)	(0.0267)	(0.0342)	(0.0384)
Wage	0.778***	0.822***	0.843***	0.824***	0.820***	0.836***
	(0.00660)	(0.00770)	(0.00844)	(0.00769)	(0.00895)	(0.0151)
Wage dispersion	0.976	0.917		0.899	0.718	1.453
	(0.176)	(0.165)		(0.163)	(0.162)	(0.435)
College dummy		0.721***	0.716***	0.727***		
		(0.0101)	(0.0108)	(0.0102)		
Children		0.562***	0.562***	0.565***	0.556***	0.503***
		(0.00763)	(0.00768)	(0.00770)	(0.00939)	(0.0116)
Observations	951,328	951,328	950,866	951,328	699,343	251,985
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 23: Estimates of the divorce equation for women. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

Moment	Data	Model
Average probability of migration (single women)	0.77%	0.87%
Average probability of marriage (women)	9.35%	9.39%
Average probability of divorce (women)	2.98%	0.51%

Table 24: Other non targeted moments.

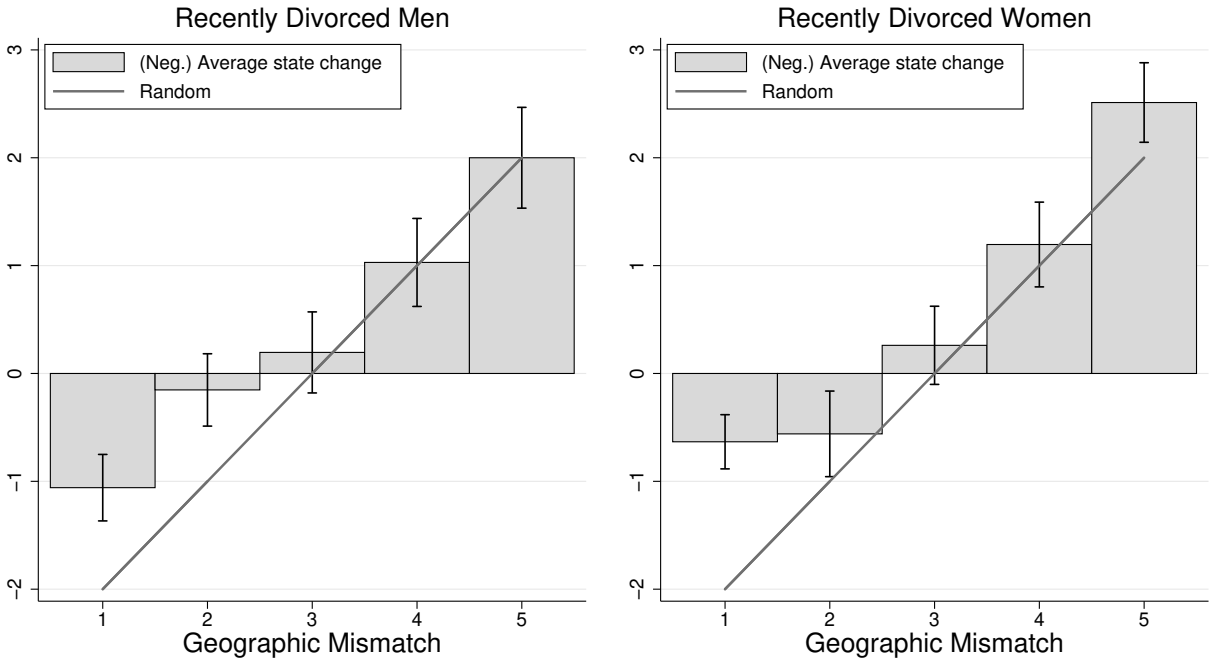

Figure 23: Average (negative) change in geographic mismatch conditional on migration for recent divorcees.

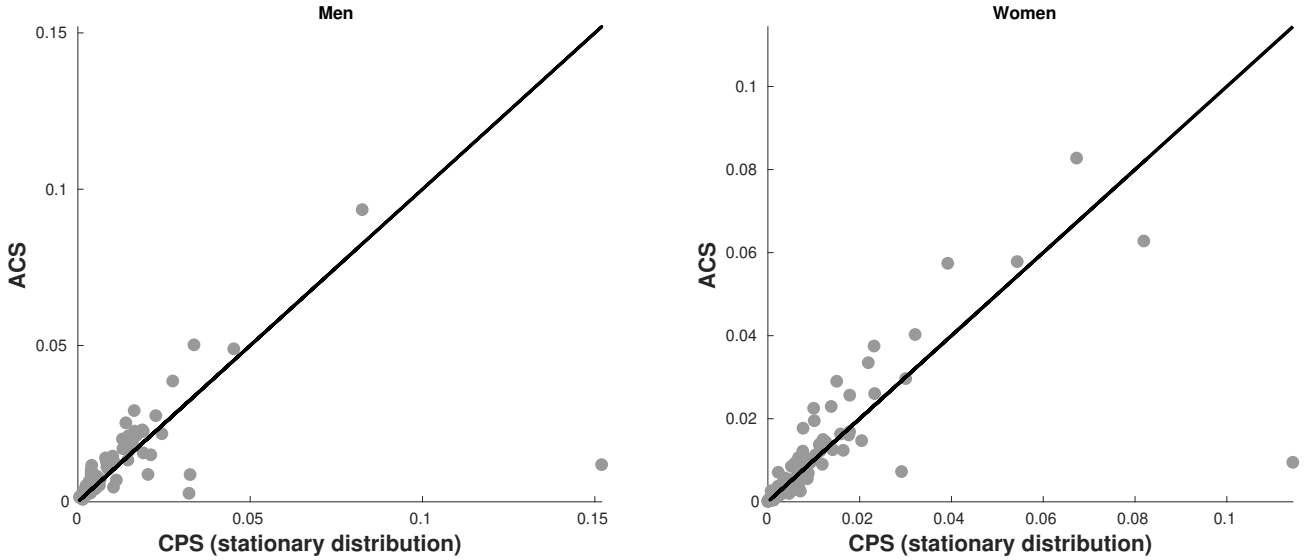

Figure 24: Comparison of the occupation shares of men and women computed from the ACS sample to the stationary distribution implied by the transition matrices obtained from the CPS.

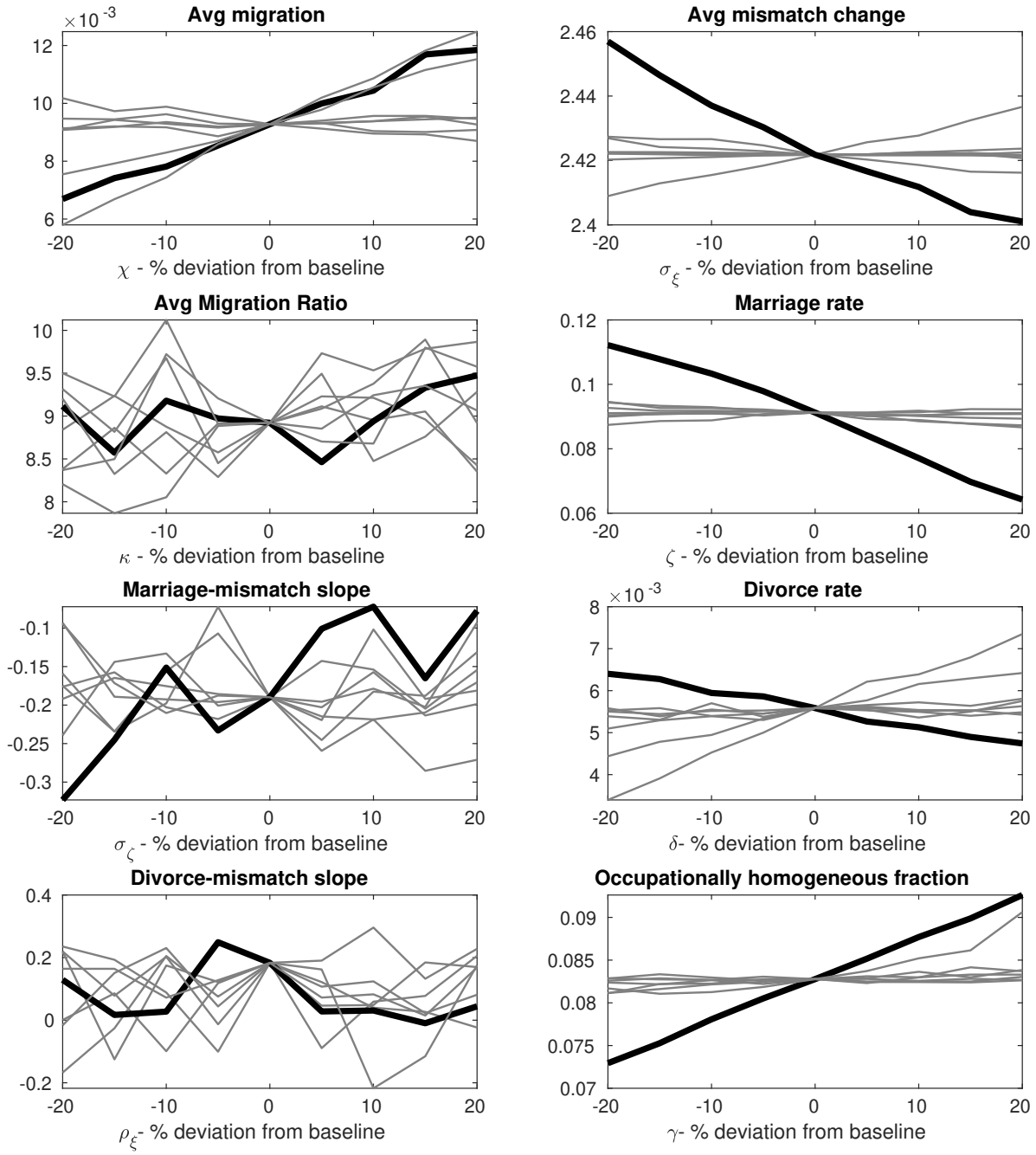

Figure 25: This graph illustrates identification by plotting the values of each identifying variable obtained by changing parameter values. Black lines correspond to the parameters identified by the variable in question, while gray lines to all the remaining parameters.

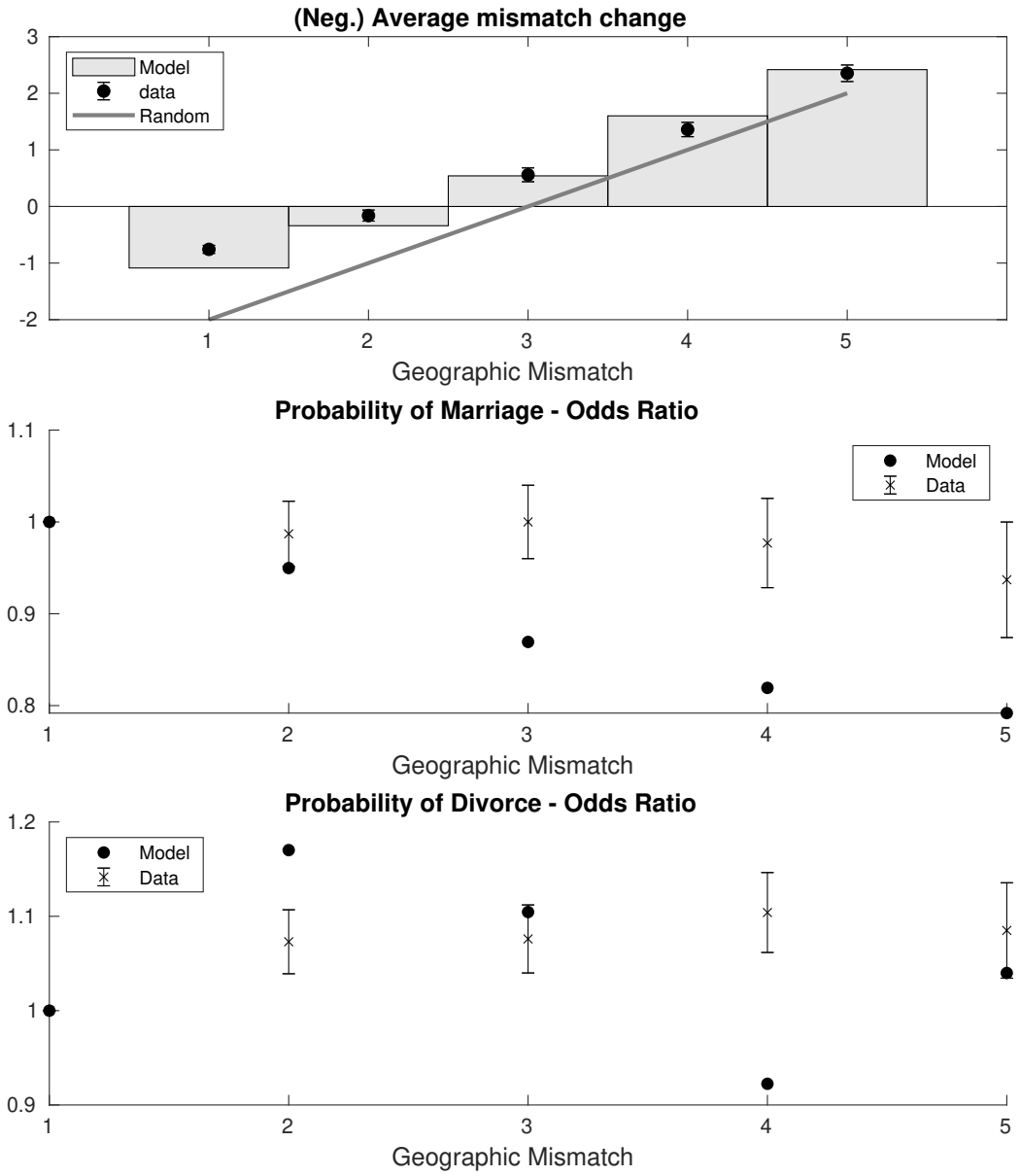

Figure 26: Estimates of the auxiliary model from real and simulated data for women. Unlike their male counterparts, these have not been targeted.

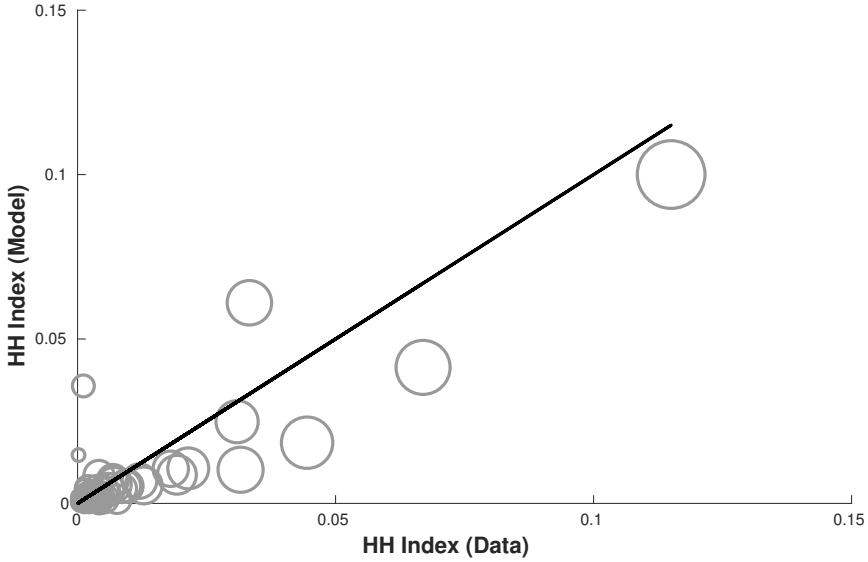

Figure 27: Goodness of fit: HH index of the geographic concentration of employment by occupation.

C Robustness: alternative occupational systems

In this appendix, we perform the empirical analysis using an alternative occupational system, namely the 3-digit system used by Dorn (2009), and report all the relevant tables and graphs. This system contains a finer definition of occupations which causes some occupations to be unobserved in some cities, implying that not all cities appear in each occupation-specific ranking.

	Yearly Probability of Migration (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.146***	1.195***	1.217***	1.248***	1.346***	0.839*
	(0.0567)	(0.0592)	(0.0627)	(0.0699)	(0.0772)	(0.0807)
Mismatch 3	1.385***	1.472***	1.500***	1.616***	1.631***	1.099
	(0.0695)	(0.0741)	(0.0799)	(0.0990)	(0.0973)	(0.101)
Mismatch 4	1.403***	1.534***	1.609***	1.700***	1.734***	1.117
	(0.0805)	(0.0884)	(0.0989)	(0.118)	(0.121)	(0.111)
Mismatch 5	2.231***	2.464***	2.577***	2.798***	3.012***	1.600***
	(0.139)	(0.155)	(0.171)	(0.218)	(0.230)	(0.171)
Wage	1.153***	1.029	0.922***	1.021	1.072*	0.970
	(0.0344)	(0.0311)	(0.0289)	(0.0320)	(0.0387)	(0.0559)
Wage dispersion	0.872	0.819		0.783	0.718	1.129
	(0.182)	(0.172)		(0.171)	(0.182)	(0.425)
Children		0.761***	0.800***	0.772***	0.663***	0.855*
		(0.0515)	(0.0547)	(0.0528)	(0.0714)	(0.0758)
College dummy		1.750***	1.408***	1.691***		
		(0.0695)	(0.0634)	(0.0674)		
Observations	453,225	453,225	443,765	437,049	262,559	190,666
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 25: Estimates of the migration equation for men using Dorn's (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

	Yearly Probability of Migration (Odds-ratios) - Women					
	All Women				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.239***	1.291***	1.279***	1.487***	1.312***	1.181
	(0.0669)	(0.0695)	(0.0718)	(0.0915)	(0.0766)	(0.164)
Mismatch 3	1.138**	1.194***	1.173**	1.490***	1.218***	1.086
	(0.0688)	(0.0720)	(0.0743)	(0.107)	(0.0803)	(0.163)
Mismatch 4	1.810***	1.920***	1.888***	2.346***	1.951***	1.784***
	(0.111)	(0.117)	(0.125)	(0.179)	(0.131)	(0.264)
Mismatch 5	2.743***	2.936***	2.924***	3.628***	3.165***	2.173***
	(0.190)	(0.204)	(0.213)	(0.316)	(0.239)	(0.367)
Wage	1.169***	0.973	0.902***	0.958	0.969	1.020
	(0.0405)	(0.0355)	(0.0360)	(0.0360)	(0.0393)	(0.0892)
Wage dispersion	1.783**	1.328		1.380	1.862**	0.165**
	(0.446)	(0.343)		(0.374)	(0.514)	(0.115)
Children		0.593***	0.626***	0.593***	0.523***	0.735***
		(0.0359)	(0.0384)	(0.0363)	(0.0418)	(0.0752)
College dummy		1.934***	1.768***	1.867***		
		(0.107)	(0.106)	(0.105)		
Observations	352,007	352,007	344,655	340,019	249,211	102,796
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 26: Estimates of the migration equation for women using Dorn (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

	Yearly Probability of Marriage (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Type 2	0.935***	0.920***	0.933***	0.934***	0.925***	0.895***
	(0.0166)	(0.0165)	(0.0171)	(0.0181)	(0.0189)	(0.0340)
Type 3	0.913***	0.891***	0.899***	0.909***	0.911***	0.838***
	(0.0182)	(0.0179)	(0.0186)	(0.0205)	(0.0214)	(0.0337)
Type 4	0.879***	0.857***	0.874***	0.874***	0.863***	0.851***
	(0.0210)	(0.0208)	(0.0219)	(0.0233)	(0.0255)	(0.0369)
Type 5	0.830***	0.790***	0.798***	0.814***	0.835***	0.729***
	(0.0271)	(0.0263)	(0.0272)	(0.0288)	(0.0350)	(0.0408)
City size	0.992	0.999	0.996	1.003	1.001	1.003
	(0.00633)	(0.00643)	(0.00656)	(0.00753)	(0.00773)	(0.0121)
Sex ratio	0.0919***	0.0806***	0.118***	0.576	0.101***	0.0616***
	(0.0215)	(0.0191)	(0.0279)	(0.217)	(0.0276)	(0.0293)
Wage	1.272***	1.252***	1.173***	1.253***	1.230***	1.365***
	(0.0134)	(0.0139)	(0.0141)	(0.0140)	(0.0156)	(0.0316)
Wage dispersion	1.005	1.060		1.075	1.256**	0.657**
	(0.0840)	(0.0896)		(0.0914)	(0.122)	(0.114)
Children		2.832***	2.907***	2.859***	3.057***	2.643***
		(0.0519)	(0.0546)	(0.0526)	(0.0723)	(0.0756)
College dummy		1.574***	1.387***	1.571***		
		(0.0279)	(0.0274)	(0.0279)		
Observations	308,453	308,453	308,203	308,453	214,562	93,891
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 27: Estimates of the marriage equation for men using Dorn's (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

	Yearly Probability of Marriage (Odds-ratios) - Women				
	All Women			College	Non-College
Type 1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)
Type 2	1.027	1.030*	1.049**	1.023	1.041
	(0.0185)	(0.0185)	(0.0208)	(0.0200)	(0.0488)
Type 3	1.038*	1.039*	1.063***	1.033	1.049
	(0.0209)	(0.0210)	(0.0246)	(0.0229)	(0.0524)
Type 4	1.012	1.012	1.030	0.989	1.084
	(0.0250)	(0.0251)	(0.0285)	(0.0273)	(0.0620)
Type 5	0.982	0.982	1.010	1.007	0.900
	(0.0337)	(0.0338)	(0.0373)	(0.0389)	(0.0695)
City size	0.995	0.995	0.995	0.987*	1.027*
	(0.00675)	(0.00680)	(0.00785)	(0.00750)	(0.0163)
Sex ratio	0.809	0.888	6.408***	0.716	2.899**
	(0.179)	(0.196)	(2.377)	(0.175)	(1.531)
Wage	1.317***	1.240***	1.243***	1.258***	1.232***
	(0.0154)	(0.0155)	(0.0156)	(0.0175)	(0.0362)
Wage dispersion	0.722***	0.706***	0.718***	0.671***	1.031
	(0.0718)	(0.0709)	(0.0726)	(0.0739)	(0.253)
Children		1.144***	1.149***	1.261***	0.928**
		(0.0211)	(0.0213)	(0.0260)	(0.0332)
College dummy		1.465***	1.464***		
		(0.0313)	(0.0313)		
Observations	275,608	275,608	275,608	214,525	61,083
Year dummies	NO	YES	YES	YES	YES
Occ. FE	NO	NO	NO	NO	NO
State FE	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 28: Estimates of the marriage equation for women using Dorn's (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation. The regression with occupation fixed effects is not available due to numerical difficulties in performing the estimation.

	Probability of Marriage Within Occupation (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.242***	1.249***	1.032	1.385***	1.190**	1.695***
	(0.0856)	(0.0863)	(0.0769)	(0.113)	(0.0880)	(0.344)
Mismatch 3	1.631***	1.653***	1.199**	1.914***	1.586***	2.229***
	(0.120)	(0.122)	(0.0990)	(0.175)	(0.127)	(0.469)
Mismatch 4	2.178***	2.220***	1.332***	2.557***	2.291***	2.209***
	(0.200)	(0.205)	(0.143)	(0.269)	(0.230)	(0.525)
Mismatch 5	2.071***	2.156***	1.199	2.457***	2.344***	1.674*
	(0.272)	(0.284)	(0.188)	(0.356)	(0.347)	(0.500)
Other sex fraction	1.419***	1.418***	1.997***	1.423***	1.374***	1.697***
	(0.0328)	(0.0332)	(0.0831)	(0.0338)	(0.0349)	(0.107)
College		0.492***	0.392***	0.496***		
		(0.0859)	(0.0756)	(0.0869)		
College, spouse		0.492***	0.447***	0.497***	1.703***	0.494***
		(0.0716)	(0.0693)	(0.0726)	(0.251)	(0.0758)
College interaction		3.471***	2.817***	3.461***		
		(0.704)	(0.617)	(0.705)		
Wage	0.333***	0.330***	0.376***	0.337***	0.308***	0.417*
	(0.0919)	(0.0891)	(0.115)	(0.0916)	(0.111)	(0.214)
Wage, spouse	0.455**	0.454***	0.537*	0.459**	0.450**	0.379*
	(0.140)	(0.137)	(0.186)	(0.139)	(0.182)	(0.196)
Wage interaction	1.444***	1.439***	1.359***	1.430***	1.455***	1.451**
	(0.130)	(0.127)	(0.137)	(0.127)	(0.169)	(0.247)
Observations	17,453	17,453	16,860	17,442	14,640	2,813
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 29: Estimates of the equation for marrying within occupation for men using Dorn's (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

	Probability of Marriage Within Occupation (Odds-ratios) - Women					
	All Women				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.123*	1.144*	1.096	1.157*	1.185**	0.771
	(0.0771)	(0.0791)	(0.0819)	(0.0945)	(0.0851)	(0.191)
Mismatch 3	1.433***	1.474***	1.340***	1.525***	1.481***	1.345
	(0.105)	(0.108)	(0.103)	(0.140)	(0.115)	(0.311)
Mismatch 4	1.827***	1.862***	1.719***	1.938***	1.812***	1.945***
	(0.165)	(0.169)	(0.170)	(0.212)	(0.178)	(0.490)
Mismatch 5	1.834***	1.887***	1.699***	1.942***	1.892***	1.550
	(0.240)	(0.248)	(0.246)	(0.283)	(0.278)	(0.485)
Other sex fraction	1.461***	1.461***	1.589***	1.457***	1.437***	1.730***
	(0.0325)	(0.0328)	(0.0386)	(0.0334)	(0.0334)	(0.147)
College		0.578***	0.519***	0.588***		
		(0.0834)	(0.0807)	(0.0854)		
College, spouse		0.837	0.448***	0.856	2.310***	0.864
		(0.144)	(0.0886)	(0.147)	(0.239)	(0.155)
College interaction		2.783***	2.250***	2.727***		
		(0.555)	(0.494)	(0.546)		
Wage	0.439***	0.399***	0.392**	0.397***	0.414**	0.427**
	(0.130)	(0.119)	(0.145)	(0.120)	(0.160)	(0.165)
Wage, spouse	0.340***	0.306***	0.303***	0.304***	0.324***	0.271***
	(0.0882)	(0.0791)	(0.0976)	(0.0799)	(0.110)	(0.107)
Wage interaction	1.448***	1.469***	1.456***	1.470***	1.452***	1.465***
	(0.125)	(0.126)	(0.156)	(0.128)	(0.159)	(0.207)
Observations	18,633	18,633	16,321	18,624	16,231	2,402
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 30: Estimates of the equation for marrying within occupation for women using Dorn's (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

	Yearly Probability of Divorce (Odds-ratios) - Men					
	All Men				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.097***	1.104***	1.101***	1.064***	1.113***	1.079***
	(0.0190)	(0.0194)	(0.0197)	(0.0204)	(0.0247)	(0.0313)
Mismatch 3	1.135***	1.132***	1.119***	1.070***	1.153***	1.087***
	(0.0207)	(0.0210)	(0.0212)	(0.0224)	(0.0276)	(0.0321)
Mismatch 4	1.192***	1.184***	1.163***	1.105***	1.234***	1.100***
	(0.0241)	(0.0244)	(0.0247)	(0.0258)	(0.0343)	(0.0344)
Mismatch 5	1.157***	1.175***	1.166***	1.097***	1.194***	1.114***
	(0.0293)	(0.0302)	(0.0305)	(0.0310)	(0.0437)	(0.0411)
Wage	0.739***	0.828***	0.874***	0.829***	0.862***	0.762***
	(0.00698)	(0.00891)	(0.0103)	(0.00891)	(0.0117)	(0.0132)
Wage dispersion	0.882	0.840**		0.863*	0.846	0.837
	(0.0769)	(0.0725)		(0.0747)	(0.0987)	(0.108)
College dummy		0.693***	0.789***	0.692***		
		(0.00982)	(0.0128)	(0.00983)		
Children		0.173***	0.173***	0.173***	0.170***	0.160***
		(0.00248)	(0.00249)	(0.00250)	(0.00322)	(0.00349)
Observations	947,948	947,948	947,496	947,948	636,425	311,523
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 31: Estimates of the divorce equation for men using Dorn's (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for his occupation.

	Yearly Probability of Divorce (Odds-ratios) - Women					
	All Women				College	Non-College
Mismatch 1	1	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)
Mismatch 2	1.043***	1.034**	1.034**	0.989	1.066***	0.961
	(0.0169)	(0.0168)	(0.0171)	(0.0180)	(0.0211)	(0.0276)
Mismatch 3	1.061***	1.056***	1.064***	0.978	1.081***	0.993
	(0.0182)	(0.0182)	(0.0188)	(0.0196)	(0.0228)	(0.0297)
Mismatch 4	1.056***	1.052***	1.061***	0.974	1.089***	0.975
	(0.0205)	(0.0205)	(0.0213)	(0.0221)	(0.0264)	(0.0323)
Mismatch 5	1.076***	1.073***	1.093***	0.985	1.118***	0.991
	(0.0260)	(0.0260)	(0.0271)	(0.0266)	(0.0344)	(0.0391)
Wage	0.776***	0.821***	0.848***	0.822***	0.819***	0.834***
	(0.00657)	(0.00768)	(0.00892)	(0.00767)	(0.00890)	(0.0151)
Wage dispersion	0.834*	0.825**		0.849*	0.691***	1.204
	(0.0824)	(0.0807)		(0.0834)	(0.0859)	(0.191)
College dummy		0.720***	0.826***	0.726***		
		(0.0100)	(0.0127)	(0.0101)		
Children		0.561***	0.573***	0.565***	0.556***	0.503***
		(0.00762)	(0.00779)	(0.00769)	(0.00937)	(0.0116)
Observations	951,294	951,294	950,831	951,294	699,318	251,976
Year dummies	NO	YES	YES	YES	YES	YES
Occ. FE	NO	NO	YES	NO	NO	NO
State FE	NO	NO	NO	YES	NO	NO

Robust seeform in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 32: Estimates of the divorce equation for women using Dorn's (2009) occupational system. The table reports the odds-ratios relative to an individual who is geographically well matched (level 1), namely an individual who lives in the highest paying quintile of cities for her occupation.

D Stationary Distribution

In this appendix, we report the equations defining the stationary distribution of the model. The equations for men and women are specular and we report only the ones for women.

Given the complexity of the transition dynamics, it is convenient to split the problem and to report three sets of equations: the first describes the distribution dynamics from the inception of the period to after the marriage market phase; the second describes the transition from the latter to the end of the period but before the preference shocks and death; the third describes the dynamics of the preference and occupation shocks and death.

Distributions after the marriage market

Let $\hat{\mu}_{c,g,x_g}$ be the mass of singles of gender g of state x_g living in city c after the marriage market phase. This fraction of the population will be equal to the fraction of singles that do not marry, namely the sum of those who do not receive a match in the marriage market and those who do match but do not marry. For women, this is given by,

$$\hat{\mu}_{c,f,x_f} = \left(1 - \int_{x_m} g_{f,c,x_f,x_m} dx_m\right) \mu_{c,f,x_f} + \int_{\zeta} \int_{x_m} g_{f,c,x_f,x_m} m\left(c, x_m, x_f, \zeta\right) \mu_{c,f,x_f} dx_m df\left(\zeta\right). \tag{29}$$

As for married households, the only thing that happens in this initial phase is the realization of the marriage quality shock. Thus the fraction of couples of type (x_m, x_f, ζ) living in c is given by

$$\hat{\tilde{\mu}}_{c,x_m,x_f,\zeta} = \int_{\zeta'} \tilde{\mu}_{c,x_m,x_f,\zeta'} df_{\zeta}\left(\zeta|\zeta'\right). \tag{30}$$

Distributions after the labor market

The mass of singles of gender g of type x_g living in city c after the labor market, $\mathring{\mu}_{c,g,x_g}$, is given by the sum of singles of the same type who were already living in c and did not move (either because they did not have the chance or did not want to) and those who were living elsewhere and moved to c. In addition, some married individuals might divorce and become single. As for the previously single, divorced individuals might be living in c because they were already living there and did not move (either by choice or by lack of opportunity), or because they moved there from elsewhere. These dynamics are described by the following

equation:

$$\mathring{\mu}_{c,f,x_f} = \left[(1-\chi) + \chi \sum_{c'|c'\neq c} \theta\left(c'|c\right) \int_{\tilde{xi}_f} \left(1 - t\left(c, x_f, c', \tilde{xi}_f\right)\right) df_\xi\left(\tilde{xi}_f\right)\right] \hat{\mu}_{c,f,x_f}$$
$$+ \chi \sum_{c|c\neq c'} \theta\left(c|c'\right) \int_{\xi'_f} t\left(c', x'_f, c, \xi_f\right) f_\xi\left(\xi_f\right) d\hat{\mu}_{c',f,x'_f}$$
$$+ (1-\chi) \int_\zeta d_{nof}\left(c, x_m, x_f, \zeta\right) d\hat{\tilde{\mu}}_{c,x_m,x_f,\zeta}$$
$$+ \chi \sum_{c|c\neq c'} \theta\left(c|c'\right) \int_{\xi'_m,\xi'_f,\zeta} d_{of}\left(c', x'_m, x'_f, \zeta, c, \xi_m, \xi_f\right) t_{out,f}\left(c', x'_f, c, \xi_f\right) f_\xi\left(\xi_m\right) f_\xi\left(\xi_f\right) d\hat{\tilde{\mu}}_{c',x'_m,x'_f,\zeta}. \tag{31}$$

Similarly, the mass of type-(x_m, x_f, ζ) married households living in c after the labor market is given by the sum of those couples who did not divorce and did not move (either by choice or by lack of opportunity) and those who moved to c from elsewhere, namely

$$\mathring{\tilde{\mu}}_{c,x_m,x_f,\zeta} = \Bigg[(1-\chi)\left(1 - d_{nof}\left(c, x_m, x_f, \zeta\right)\right) + \chi \sum_{c'|c'\neq c} \theta\left(c'|c\right)$$
$$\int_{\tilde{\xi}_m} \int_{\tilde{\xi}_f} \left(1 - t\left(c, x_m, x_f, \zeta, c', \tilde{\xi}_m, \tilde{\xi}_f\right)\right)$$
$$\left(1 - d_{of}\left(c, x_m, x_f, \zeta, c', \tilde{\xi}_m, \tilde{\xi}_f\right)\right) df_\xi\left(\tilde{\xi}_m\right) df_\xi\left(\tilde{\xi}_f\right)\Bigg] \hat{\tilde{\mu}}_{c,x_m,x_f,\zeta} \tag{32}$$
$$+ \chi \sum_{c|c\neq c'} \theta\left(c|c'\right) \int_{\xi'_m,\xi'_f,\zeta} \left(1 - d_{of}\left(c', x'_m, x'_f, \zeta, c, \xi_m, \xi_f\right)\right)$$
$$t\left(c', x'_m, x'_f, \zeta, c, \xi_m, \xi_f\right) f_\xi\left(\xi_m\right) f_\xi\left(\xi_f\right) d\hat{\tilde{\mu}}_{c',x'_m,x'_f,\zeta}.$$

Preference shocks and death

Finally, at the end of each period a fraction $(1-\omega)$ of households dies and is replaced by an equal number of single households with types drawn from a given distribution $F_g\left(x_g\right)$, with $g = m, f$ and probability distribution $f_g\left(x_g\right)$. Those who survive receive a preference shock and potentially change occupation. The mass of single women of each type will then be given by

$$\mu_{c,f,x_f} = \omega \sum_{j'_f} \pi\left(j_f|j'_f\right) \int_{\xi'_f} \mathring{\mu}_{c,f,x'_f} df_\xi\left(\xi_f|\xi'_f\right) + (1-\omega)\, df_f\left(x_f\right) \left[\int_{x_f} d\mathring{\mu}_{c,f,x_f} + \int_{x_m,x_f,\zeta} d\mathring{\tilde{\mu}}_{c,x_m,x_f,\zeta}\right]. \tag{33}$$

where $f_\xi\left(\xi_f|\xi'_f\right)$ is the conditional distribution of ξ. For married households we have

$$\tilde{\mu}_{c,x_m,x_f,\zeta} = \omega \int_{\xi'_m} \int_{\xi'_f} \mathring{\tilde{\mu}}_{c,x'_m,x'_f,\zeta} df_\xi\left(\xi_m|\xi'_m\right) df_\xi\left(\xi_f|\xi'_f\right) \tag{34}$$

E Existence of the Equilibrium

In appendix F, we construct an *update function* Ψ that has a fixed point if and only if there exists a stationary equilibrium. This function takes as argument an element from the set of feasible distributions of singles $\mathcal{D}$, computes every endogenous object in the model, and returns an updated distribution of singles from the set of feasible distributions. To prove the existence of an equilibrium, all we need is to prove that Ψ and $\mathcal{D}$ satisfy the conditions stated in Brouwer's theorem, namely that $\mathcal{D}$ is compact and convex and that Ψ is a continuous mapping from $\mathcal{D}$ to itself.

$\mathcal{D}$ is compact and convex. The stated properties of the set $\mathcal{D}$ follow trivially from the definition of the set itself. Let $\boldsymbol{\mu}$ be a vector of values μ_{c,g,x_g} for all c, g and x_g describing a distribution of single males and females over the state space. Then, the set of feasible distributions, $\mathcal{D}$, is defined as

$$\mathcal{D} = \left\{ \boldsymbol{\mu} | \sum_c \sum_{x_f} \mu_{c,f,x_f} \leq F; \sum_c \sum_{x_m} \mu_{c,m,x_m} \leq M \right\}. \tag{35}$$

Compactness and convexity of this set are thus trivially satisfied.

Ψ is a continuous mapping from $\mathcal{D}$ to itself. The mapping Ψ is described in detail in appendix F. We are assured that for each $\boldsymbol{\mu} \in \mathcal{D}$, $\boldsymbol{\mu'} = \Psi\left(\boldsymbol{\mu}\right) \in \mathcal{D}$ simply because the mass of males M and females F is exogenous to the model and kept constant by construction. The continuity of Ψ follows from the continuity of the bliss (ζ) and preference (ξ) shocks and the continuity of the matching function $M\left(\cdot,\cdot\right)$. Intuitively, the continuity of the shocks, ensures that there are no discrete "jumps" in the flows of workers across states as the value functions associated with each of the discrete choices that agents make change. Moreover, these value functions are continuous in the matching probabilities that are continuous in the distribution $\boldsymbol{\mu}$ because of the continuity of M. It follows that all the value functions in the model are continuous in $\boldsymbol{\mu}$. Summing up, continuous changes in $\boldsymbol{\mu}$ cause continuous changes in the value functions, that cause continuous changes in the flows of workers across states and thus continuous changes in $\boldsymbol{\mu'}$.

F Model Solution

This appendix describes the algorithm employed for the numerical solution of the model.

Step 1: Guess initial feasible[56] distributions of singles, μ_{c,m,x_m} and μ_{c,f,x_f}.

Step 2: Given the guessed distributions, compute the matching probabilities using equation (8).

Step 3: Using the probabilities computed above, solve the fixed-point problem described by equations (7), (10), (12) and (21). The solution to this problem is computed through value function iteration. As a byproduct of the process, we obtain the policy functions for marriage, divorce and migration.

Step 4: Using the policy functions obtained in the previous step, obtain a new guess for the distribution of singles using equations(29) through (34).

Step 5: Evaluate the distance between the initial guess and the new guess. If it is less than the specified tolerance then terminate. Otherwise, repeat from step 2 using the new guess as the initial guess.

Notice that the algorithm outlined above describes a mapping $\Psi : \mathcal{D} \to \mathcal{D}$ from the set of feasible distributions of singles $\mathcal{D}$ to itself. The stationary equilibrium then can be defined as the solution to the fixed point problem described by $\Psi(\mathcal{D}) = \mathcal{D}$.

G Dealing with the curse of dimensionality

In this appendix, we describe the process used to construct clusters of cities. Such clusters are used to reduce the computational power needed to numerically compute the equilibrium of the model. Since the only exogenous driver of the migration and marriage patterns in the model is the wage variation across cities and occupations, this process is aimed at clustering cities with similar wage distributions together.

To perform the clustering, we employ a process that is a combination of two unsupervised machine learning techniques. First, we perform a permanent component analysis (PCA) on the city-occupation wage premia. Then, we perform k-means clustering on a subset of the principal components computed in the previous stage. Performing k-means clustering on

[56] The distributions have to satisfy $\sum_c \int_{x_f} d\mu_{c,f,x_f} = F$ and $\sum_c \int_{x_f} d\mu_{c,m,x_m} = M$.

a subset of permanent components is not a novelty in the scientific literature (Ibes, 2015). Although there is no evidence that this two-step procedure improves the quality of the resulting clustering over a direct application of the k-means algorithm (Yeung and Ruzzo, 2001), the dimensionality reduction achieved with the PCA is extremely helpful because of the workings of the k-means algorithm. The k-means algorithm is, in fact, initialized randomly drawing initial centroids for the clusters. The algorithm then updates these centroids based on some measure of distance (Euclidean distance in this case). Issues arise when there is a high dimensionality, i.e. if the vector of characteristics on which the clustering is performed is too big. In this case, the algorithm tends to produce results that are highly dependent on the starting point. One way to deal with this is to try with different initial points and then pick the best resulting clustering. Nevertheless, with very high dimensionality, this is almost akin to a random search of the best clustering. An additional advantage of using a subset of permanent components is that it reduces the impact of the noise coming from the estimation errors from the wage regression.

To perform PCA, we consider each city as one observation and the set of occupation fixed effects (the estimates of $\alpha_{c,o}$ from equation 1) as the covariates. Figure 28 shows the share of variation explained by the first 30 principal components. The first component explains more than 30% of the variation. The explanatory power of the second component drops by almost 25 percentage points to just above 5%, with the successive components slowly losing explanatory power.

In the second step, we perform k-means on the 5 principal components. This choice is rather arbitrary as there is no objective way nor any rule of thumb. The 5 components together explain 49% of the total variation.

One of the shortcomings of k-means is that the algorithm does not identify the optimal number of clusters. The desired number of clusters has to be provided by the user. A simple and popular solution consists of inspecting the dendrogram produced using hierarchical clustering to see if it suggests a particular number of clusters. Nevertheless, this approach is also subjective. An alternative popular way is the elbow method that consists of a graphical analysis of the total intra-cluster variation as a function of the number of clusters.[57] The optimal number of clusters is identified by an "elbow" in the graph, i.e. that point from which the gains from increasing the number of clusters in terms of the reduction of the total intra-cluster variation i becomes small.

To choose the number of clusters we adapt the idea behind the elbow method to our particular problem. As discussed in the main text, the most important aspect of the wage distribution that we want to preserve is the covariance structure of the occupation-city

[57] As the number of clusters increases the intra-cluster variation is mechanically reduced.

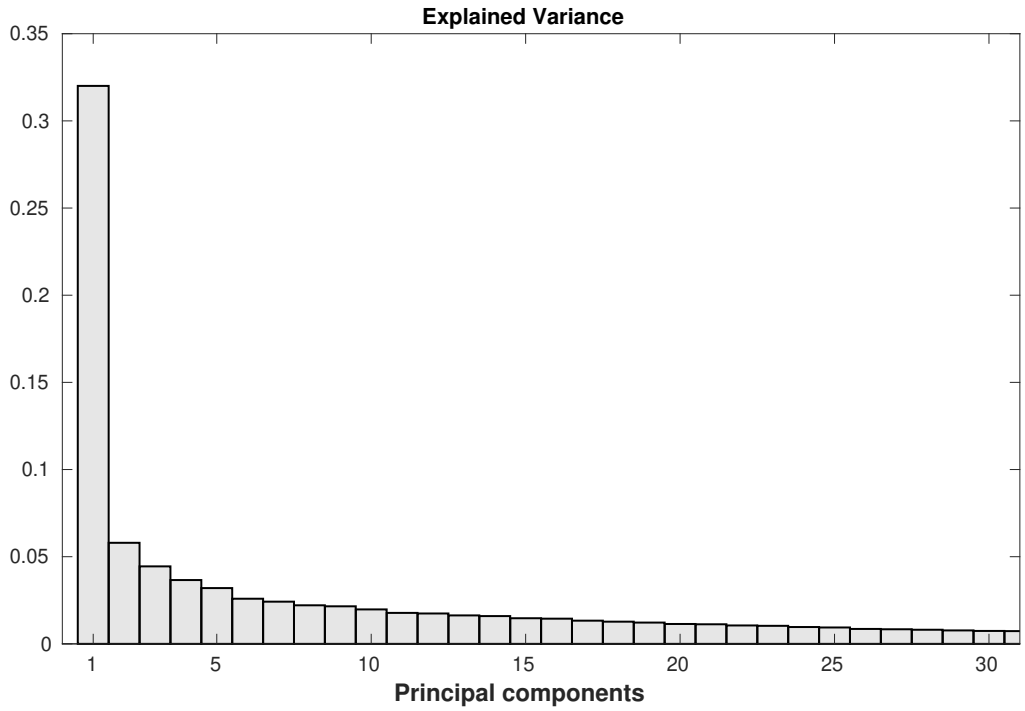

Figure 28: Explanatory power of the first 30 principal components.

premia. With this goal in mind, we perform k-means[58] varying the number of clusters from 5 to 150. For each clustering we estimate the wage regression (1) and compute the wage covariance structure.[59] Next, we compare the estimated covariance structure to the original one (i.e. the one obtained with actual cities). In practice, we compute the Euclidean distance between the covariance vectors from the cluster and the original data. The computed distances as a function of the number of clusters are shown in figure 29

As expected, the function is downward sloping as adding more clusters makes it easier to match the actual wage covariance structure. Nevertheless, the gains from adding more clusters are drastically reduced after 20/30 clusters. Given this, we deemed 25 clusters to be a fair compromise between accuracy and the need for speed. Table 33 shows the composition of the 25 clusters used in the paper.

[58]For each call of k-means, we run the clustering algorithm 10000 times for different initial points and select the clustering that returns the lowest total intra-cluster variation.

[59]Each occupation is treated as an observation.

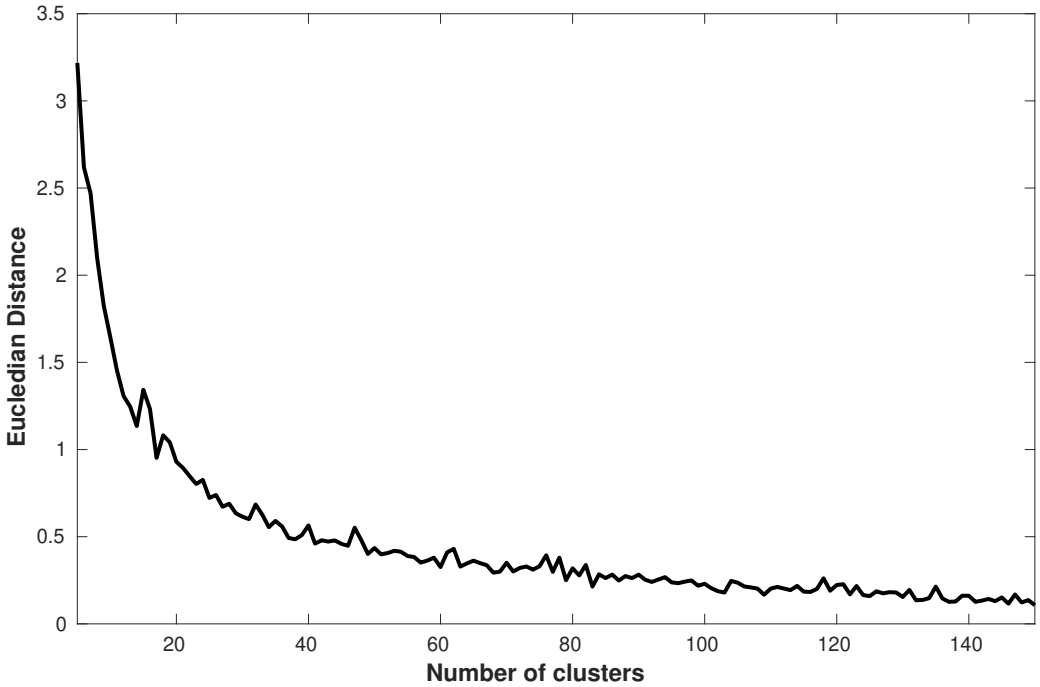

Figure 29: Euclidean distance between the actual and clustered covariance vectors as a function of the number of clusters.

Table 33: Composition of the 25 clusters.

Cluster	MSA
1	ann arbor, mi bakersfield, ca cleveland-elyria, oh columbus, oh denver-aurora-lakewood, co fresno, ca indianapolis-carmel-anderson, in iowa city, ia lancaster, pa new orleans-metairie, la phoenix-mesa-scottsdale, az portland-vancouver-hillsboro, or-wa reno, nv st. louis, mo-il virginia beach-norfolk-newport news, va-nc york-hanover, pa
2	bridgeport-stamford-norwalk, ct new york-newark-jersey city, ny-nj-pa san francisco-oakland-hayward, ca san jose-sunnyvale-santa clara, ca trenton, nj washington-arlington-alexandria, dc-va-md-wv
3	los angeles-long beach-anaheim, ca modesto, ca oxnard-thousand oaks-ventura, ca sacramento–roseville–arden-arcade, ca

	salinas, ca san diego-carlsbad, ca santa maria-santa barbara, ca
4	amarillo, tx boise city, id college station-bryan, tx deltona-daytona beach-ormond beach, fl gainesville, fl lakeland-winter haven, fl lynchburg, va santa fe, nm yuma, az
5	burlington-south burlington, vt cincinnati, oh-ky-in harrisburg-carlisle, pa kansas city, mo-ks pittsburgh, pa portland-south portland, me rochester, ny salt lake city, ut springfield, il syracuse, ny
6	albany-schenectady-troy, ny allentown-bethlehem-easton, pa-nj bloomington, il detroit-warren-dearborn, mi manchester-nashua, nh milwaukee-waukesha-west allis, wi minneapolis-st. paul-bloomington, mn-wi providence-warwick, ri-ma racine, wi riverside-san bernardino-ontario, ca springfield, ma
7	baton rouge, la beaumont-port arthur, tx dover, de houma-thibodaux, la monroe, mi yuba city, ca
8	akron, oh bend-redmond, or chico, ca decatur, il eugene, or jackson, mi lansing-east lansing, mi mobile, al niles-benton harbor, mi port st. lucie, fl redding, ca scranton–wilkes-barre–hazleton, pa toledo, oh visalia-porterville, ca yakima, wa
9	anchorage, ak barnstable town, ma bremerton-silverdale, wa

	las vegas-henderson-paradise, nv napa, ca san luis obispo-paso robles-arroyo grande, ca stockton-lodi, ca urban honolulu, hi vallejo-fairfield, ca
10	colorado springs, co daphne-fairhope-foley, al dayton, oh flagstaff, az fort collins, co grand rapids-wyoming, mi jackson, ms janesville-beloit, wi kalamazoo-portage, mi ogden-clearfield, ut oklahoma city, ok omaha-council bluffs, ne-ia provo-orem, ut saginaw, mi shreveport-bossier city, la spokane-spokane valley, wa
11	wichita falls, tx
12	atlantic city-hammonton, nj baltimore-columbia-towson, md boston-cambridge-newton, ma-nh chicago-naperville-elgin, il-in-wi hartford-west hartford-east hartford, ct new haven-milford, ct norwich-new london, ct philadelphia-camden-wilmington, pa-nj-de-md santa cruz-watsonville, ca santa rosa, ca seattle-tacoma-bellevue, wa worcester, ma-ct
13	mansfield, oh medford, or
14	goldsboro, nc
15	albuquerque, nm augusta-richmond county, ga-sc chattanooga, tn-ga fayetteville, nc greensboro-high point, nc greenville-anderson-mauldin, sc knoxville, tn lubbock, tx pensacola-ferry pass-brent, fl san antonio-new braunfels, tx spartanburg, sc tyler, tx waco, tx wichita, ks winston-salem, nc
16	birmingham-hoover, al fayetteville-springdale-rogers, ar-mo jacksonville, fl lafayette, la

	little rock-north little rock-conway, ar louisville/jefferson county, ky-in memphis, tn-ms-ar nashville-davidson–murfreesboro–franklin, tn
17	cape coral-fort myers, fl gainesville, ga hilton head island-bluffton-beaufort, sc miami-fort lauderdale-west palm beach, fl naples-immokalee-marco island, fl north port-sarasota-bradenton, fl orlando-kissimmee-sanford, fl palm bay-melbourne-titusville, fl tampa-st. petersburg-clearwater, fl tucson, az
18	burlington, nc gulfport-biloxi-pascagoula, ms huntsville, al montgomery, al ocala, fl springfield, oh
19	east stroudsburg, pa glens falls, ny la crosse-onalaska, wi-mn rockford, il sheboygan, wi
20	bangor, me blacksburg-christiansburg-radford, va owensboro, ky prescott, az
21	binghamton, ny eau claire, wi elkhart-goshen, in erie, pa johnstown, pa lewiston-auburn, me lincoln, ne michigan city-la porte, in pueblo, co state college, pa utica-rome, ny wausau, wi youngstown-warren-boardman, oh-pa
22	brownsville-harlingen, tx clarksville, tn-ky el paso, tx hickory-lenoir-morganton, nc joplin, mo mcallen-edinburg-mission, tx
23	bellingham, wa buffalo-cheektowaga-niagara falls, ny champaign-urbana, il kankakee, il lebanon, pa merced, ca olympia-tumwater, wa reading, pa
24	atlanta-sandy springs-roswell, ga

	austin-round rock, tx charleston-north charleston, sc charlotte-concord-gastonia, nc-sc dallas-fort worth-arlington, tx houston-the woodlands-sugar land, tx raleigh, nc richmond, va
25	canton-massillon, oh corpus christi, tx decatur, al fort wayne, in springfield, mo

H The gender wage gap

The exogenous gender gap used in the model has to be intended as an ex-ante income gap reflecting factors that are orthogonal to marriage and divorce (e.g. discrimination). This value is, in fact, obtained from the estimation of equation (1) which controls, among others, for marital status and city fixed effects.

The interaction between marriage and divorce induces an additional income gap that is generated from the different migration patterns of men and women. To get a sense of the ex-post gender wage gap, we can compute the log difference between the average wage of men and the average wage of women from both the model and the data. For the latter, we compute the gap using the residuals from a regression of log wages over education and year dummies and a quadratic function of potential work experience. This adjustment is necessary to make the quantities from the data comparable to that of the model since the latter does not present lifecycle effects nor heterogeneity in education. Reassuringly, the empirical and modeled gender gaps are very close, the former being equal to -26.1% and the latter to -26.2%.

I Computing bounds for the change in total earnings

This appendix explains the details of calculations performed to compute the boundaries for the change in total labor earning reported in table 10.

The lower bound

To determine the lower bound for the change of total earnings, we use the demand elasticities obtained from the estimated production function from Alonzo and Gallipoli (2020). In that

paper, the labor demand side of the economy is characterized by a continuum of firms operating in monopolistic competition that demand single occupations to produce intermediate goods that are subsequently combined by a representative firm producing the consumption good. This production structure produces the inverse demand function for occupation j'

$$w_{i\hat{j}} = \rho A \alpha_{\hat{j}} \beta_{\hat{j}} \left[\sum_{\hat{j}'} \alpha_{\hat{j}'} \left(\sum_{i} \beta_{i'\hat{j}'} L_{i'\hat{j}'} \right)^{\rho} \right]^{\frac{1-\rho}{\rho}} \left(\sum_{i'} \beta_{i'\hat{j}} L_{i'\hat{j}} \right)^{\rho-1} \tag{36}$$

The corresponding elasticity of wages to labor is

$$\begin{aligned}
\epsilon_{i\hat{j}} =& \frac{\partial w_{i\hat{j}}/w_{i\hat{j}}}{\partial L_{i\hat{j}}/L_{i\hat{j}}} \\
=& A \alpha_{\hat{j}} \beta_{i\hat{j}}^2 \rho (1-\rho) \left(\sum_{i'} \beta_{i'\hat{j}} L_{i'\hat{j}} \right)^{\rho-2} \left[\sum_{\hat{j}'} \alpha_{\hat{j}'} \left(\sum_{i'} \beta_{i'\hat{j}'} L_{i'\hat{j}'} \right)^{\rho} \right]^{\frac{1-\rho}{\rho}} \left[\frac{\alpha_{\hat{j}} \left(\sum_{i'} \beta_{i'\hat{j}} L_{i'\hat{j}} \right)^{\rho}}{\sum_{\hat{j}'} \alpha_{\hat{j}'} \left(\sum_{i'} \beta_{i'\hat{j}'} L_{i'\hat{j}'} \right)^{\rho}} - 1 \right] \frac{L_{i\hat{j}}}{w_{i\hat{j}}} \\
=& \beta_{i\hat{j}} \rho (1-\rho) \left(\sum_{i'} \beta_{i'\hat{j}} L_{i'\hat{j}} \right)^{-1} \left[\frac{\alpha_{\hat{j}} \left(\sum_{i'} \beta_{i'\hat{j}} L_{i'\hat{j}} \right)^{\rho}}{\sum_{\hat{j}'} \alpha_{\hat{j}'} \left(\sum_{i'} \beta_{i'\hat{j}'} L_{i'\hat{j}'} \right)^{\rho}} - 1 \right] L_{i\hat{j}} \\
=& \beta_{i\hat{j}} \rho (1-\rho) \frac{L_{i\hat{j}}}{\left(\sum_{i'} \beta_{i'\hat{j}} L_{i'\hat{j}} \right)} \left[\frac{\alpha_{\hat{j}} \left(\sum_{i'} \beta_{i'\hat{j}} L_{i'\hat{j}} \right)^{\rho}}{\sum_{\hat{j}'} \alpha_{\hat{j}'} \left(\sum_{i'} \beta_{i'\hat{j}'} L_{i'\hat{j}'} \right)^{\rho}} - 1 \right]
\end{aligned}$$

These elasticities cannot be directly applied to the framework of this paper for two reasons: (i) our occupational system is more disaggregated than in Alonzo and Gallipoli (2020) and (ii) the demographic definition is more aggregated. To overcome the first issue, we compute an employment-weighted average of these elasticities for each occupation.

$$\epsilon_{\hat{j}} = \sum_{i} \mu_{i\hat{j}} \epsilon_{i\hat{j}} \tag{37}$$

where $\mu_{i\hat{j}}$ is the share of workers in occupation $\hat{j}$ from in demographics i. To deal with the second issue, we simply assume that the elasticities are constant for all the occupations composing each aggregated occupation used in Alonzo and Gallipoli (2020). In other words, let $\Xi_{\hat{j}}$ be the set of occupations j (from the occupational system used in this paper) that are contained in the definition of occupation $\hat{j}$ (from the occupational system used by Alonzo and Gallipoli, 2020); we assume that $\epsilon_j = \epsilon_{\hat{j}}$ for all $j \in \Xi_{\hat{j}}$.

To compute the lower bound for the change in total labor earnings, we compute the above elasticities (using the estimates for the year 2018) for each city separately in the baseline model and use them to compute the new wages in the counterfactual scenarios.

The upper bound

To compute an upper bound for total labor income in the counterfactual scenarios, we rely on the literature on agglomeration economies. In a meta-analysis of 729 elasticities of wages to city size, Melo, Graham and Noland (2009) identify an average elasticity of 3.2% with a standard deviation of 7.6%.[60] To compute the counterfactual wages, we use a conservative value for the elasticity of 18.4% (the mean plus two standard deviations).

J Occupation System

In this appendix, we report the occupation system used in the paper ("Occda"). This system is the result of a balancing exercise between two requirements that push towards different levels of aggregation. On the one hand, the nature of the analysis carried out in the paper requires as much occupational heterogeneity as possible, since the variation in the wage distributions faced by each occupation is the main source of identification. On the other hand, too much disaggregation creates issues both on the empirical and the theoretical side. Both the empirical analysis and the calibration of the model require the estimation of occupation-specific city premia for each city, that is possible only if we can observe workers for each occupation in each city. Moreover, the numerical solution of the model is computationally demanding, and the required computational power is increasing in the number of occupations, thus reducing the number of occupations helps to reduce the computational burden.

To address these challenges, we developed the occupational system described in table 34 as a partially aggregated version of David Dorn's occupational system (Dorn, 2009). The advantage of starting from his classification (called "Occ1990dd") resides in the fact that a crosswalk to the 2008-2009 ACS occupation codes is freely provided by the author. Using the tables provided by IPUMS, we extended this crosswalk to the 2010-2011 ACS and 2012-2017 ACS occupation codes,[61] and then we applied the crosswalk to our classification described in the table below. In performing this aggregation, we follow two principles. The first is to perform the minimum amount of aggregation that leads to observing each aggregated occupation in each city. The second is to combine similar occupations using as guidance data on the task content provided by Autor and Dorn (2013).

Eventually, we are left with 96 occupations. In the paper, we drop extractive occupations (code 614) since they are geographically very concentrated.

[60] See Table 2 in Melo, Graham and Noland (2009).

[61] The discrepancies across these ACS years are minimal.

Table 34: Crosswalk between Dorn (2009) Occ1990 and our Occda occupations systems

Occupation Titles	Occda Code	Occ1990dd Code	Occupation Titles (Dorn, 2009)
Chief executives, public administrators, and legislators	4	4	Chief executives, public administrators, and legislators
Financial managers	7	7	Financial managers
Human resources and labor relations managers	8	8	Human resources and labor relations managers
Managers and specialists in marketing, advert., PR	13	13	Managers and specialists in marketing, advert., PR
Managers in education and related fields	14	14	Managers in education and related fields
Managers of medicine and health occupations	15	15	Managers of medicine and health occupations
Managers and administrators, n.e.c.	22	18	Managers of properties and real estate
		19	Funeral directors
		22	Managers and administrators, n.e.c.
Accountants and auditors	23	23	Accountants and auditors
Other financial specialists	25	24	Insurance underwriters
		25	Other financial specialists
Personnel, HR, training, and labor rel. specialists	27	27	Personnel, HR, training, and labor rel. specialists
Purchasing managers, agents, and buyers, n.e.c.	33	28	Purchasing agents and buyers of farm products
		29	Buyers, wholesale and retail trade
		33	Purchasing managers, agents, and buyers, n.e.c.
Inspectors and compliance officers	36	35	Construction inspectors
		36	Inspectors and compliance officers, outside
Management support occupations	37	26	Management analysts
		34	Business and promotion agents
		37	Management support occupations
Architects and civil engineers	53	43	Architects
		53	Civil engineers
Engineers and other professionals, n.e.c.	59	44	Aerospace engineers
		45	Metallurgical and materials engineers
		47	Petroleum, mining, and geological engineers
		48	Chemical engineers
		55	Electrical engineers
		56	Industrial engineers
		57	Mechanical engineers
		59	Engineers and other professionals, n.e.c.
Computer systems analysts and computer scientists	64	64	Computer systems analysts and computer scientists
Operations and systems researchers and analysts, actuaries and mathematicians	65	65	Operations and systems researchers and analysts
		66	Actuaries
		68	Mathematicians and statisticians
Physical scientists	76	69	Physicists and astronomists
		73	Chemists
		74	Atmospheric and space scientists
		75	Geologists
		76	Physical scientists, n.e.c.
		77	Agricultural and food scientists
		78	Biological scientists
		79	Foresters and conservation scientists
Medical scientists	83	83	Medical scientists
		84	Physicians
Health and therapy occupations	89	85	Dentists
		86	Veterinarians
		87	Optometrists

		88	Podiatrists
		89	Other health and therapy occupations
Pharmacists, dietitians and other therapists	105	95	Registered nurses
		96	Pharmacists
		97	Dietitians and nutritionists
		98	Respiratory therapists
		99	Occupational therapists
		103	Physical therapists
		104	Speech therapists
		105	Therapists, n.e.c.
		106	Physicians' assistants
Subject instructors, college	154	154	Subject instructors, college
Teachers	159	155	Kindergarten and earlier school teachers
		156	Primary school teachers
		157	Secondary school teachers
		158	Special education teachers
		159	Teachers, n.e.c.
Vocational and educational counselors, librarians and archivist	164	163	Vocational and educational counselors
		164	Librarians
		165	Archivists and curators
Social scientists	169	166	Economists, market and survey researchers
		167	Psychologists
		169	Social scientists and sociologists, n.e.c.
Urban and regional planners and social workers	174	173	Urban and regional planners
		174	Social workers
Clergy and religious workers	176	176	Clergy and religious workers
Welfare service workers	177	177	Welfare service workers
Lawyers and judges	178	178	Lawyers and judges
Designers	185	185	Designers
Arts and entertainment workers	194	186	Musicians and composers
		187	Actors, directors, and producers
		188	Painters, sculptors, craft-artists, and print-makers
		189	Photographers
		193	Dancers
		194	Art/entertainment performers and related occs
Residual professionals	199	183	Writers and authors
		184	Art/entertainment performers and related occs
		195	Editors and reporters
		198	Announcers
		199	Athletes, sports instructors, and officials
Health technologists and technicians	208	203	Clinical laboratory technologies and technicians
		204	Dental hygienists
		205	Health record technologists and technicians
		206	Radiologic technologists and technicians
		207	Licensed practical nurses
		208	Health technologists and technicians, n.e.c.
Engineering technicians	214	214	Engineering technicians
Drafters and surveyors	217	217	Drafters
		218	Surveyors, cartographers, mapping scientists/techs
Programmers	229	229	Computer software developers
		233	Programmers of numerically controlled machine tools

Legal assistants and paralegals	234	234	Legal assistants and paralegals
Technicians	235	223	Biological technicians
		224	Chemical technicians
		225	Other science technicians
		226	Airplane pilots and navigators
		227	Air traffic controllers
		228	Broadcast equipment operators
		235	Technicians, n.e.c.
Sales supervisors and proprietors	243	243	Sales supervisors and proprietors
Skilled salespersons	253	253	Insurance sales occupations
		254	Real estate sales occupations
		255	Financial service sales occupations
		256	Financial service sales occupations
		258	Sales engineers
Salespersons, n.e.c.	274	274	Salespersons, n.e.c.
		275	Retail salespersons and sales clerks
		277	Door-to-door sales, street sales, and news vendors
		283	Sales demonstrators, promoters, and models
Cashiers	276	276	Cashiers
Office supervisors	303	303	Office supervisors
Computer and peripheral equipment operators	313	308	Computer and peripheral equipment operators
		313	Secretaries and stenographers
		315	Typists
		316	Interviewers, enumerators, and surveyors
Hotel and information clercks	319	317	Hotel clerks
		318	Transportation ticket and reservation agents
		319	Receptionists and other information clerks
Human resources clerks	326	326	Correspondence and order clerks
		328	Human resources clerks, excl payroll and timekeeping
Library assistants, file clerks and bookkeepers	335	329	Library assistants
		335	File clerks
		336	Records clerks
		337	Bookkeepers and accounting and auditing clerks
		338	Payroll and timekeeping clerks
Office machine operators	347	344	Billing clerks and related financial records processing
		346	Mail and paper handlers
		347	Office machine operators, n.e.c.
		348	Telephone operators
		349	Other telecom operators
Postal clerks	354	354	Postal clerks, exluding mail carriers
		355	Mail carriers for postal service
		356	Mail clerks, outside of post office
Messengers and dispatchers	357	357	Messengers
		359	Dispatchers
Stock and inventory clerks	364	364	Shipping and receiving clerks
		365	Stock and inventory clerks
		366	Meter readers
		368	Weighers, measurers, and checkers
		373	Material recording, sched., prod., plan., expediting cl.
Examiners and investigators	375	375	Insurance adjusters, examiners, and investigators

		376	Customer service reps, invest., adjusters, excl. insur.
General office clerks	379	377	Eligibility clerks for government prog., social welfare
		378	Bill and account collectors
		379	General office clerks
Administrative support jobs	389	383	Bank tellers
		384	Proofreaders
		385	Data entry keyers
		386	Statistical clerks
		387	Teacher's aides
		389	Administrative support jobs, n.e.c.
Housekeepers and laundry workers	405	405	Housekeepers, maids, butlers, and cleaners
		408	Laundry and dry cleaning workers
Police and fire fighting occupations	417	415	Supervisors of guards
		417	Fire fighting, fire prevention, and fire inspection occs
		418	Police and detectives, public service
Sheriffs, bailiffs, correctional institution officers and crossing guards	423	423	Sheriffs, bailiffs, correctional institution officers
		425	Crossing guards
Protective service, n.e.c.	427	426	Guards and police, except public service
		427	Protective service, n.e.c.
Bartenders and waiters	434	434	Bartenders
		435	Waiters and waitresses
Food preparation and service workers	439	433	Supervisors of food preparation and service
		436	Cooks
		439	Food preparation workers
		444	Miscellanious food preparation and service workers
Health assistants	445	445	Dental Assistants
		447	Health and nursing aides
Supervisors of landscaping and building service	448	448	Supervisors of cleaning and building service
		450	Superv. of landscaping, lawn service, groundskeeping
Gardeners, janitors and pest control occupations	451	451	Gardeners and groundskeepers
		453	Janitors
		455	Pest control occupations
Barbers and hairdressers	457	457	Barbers
		458	Hairdressers and cosmetologists
Recreation and Hospitality Occupations	459	459	Recreation and Hospitality Occupations
		461	Guides
		462	Ushers
		464	Baggage porters, bellhops and concierges
		466	Recreation and fitness workers
		467	Motion picture projectionists
Personal service occupations	468	468	Child care workers
		469	Personal service occupations, n.e.c
		470	Supervisors of personal service jobs, n.e.c
		471	Public transportation attendants and inspectors
		472	Animal caretakers, except farm
Farmers operators and Managers	473	473	Farmers (owners and tenants)
		475	Farm managers
Other agricultural and related occupations	479	479	Farm workers, incl. nursery farming
		488	Graders and sorters of agricultural products
		489	Inspectors of agricultural products

		496	Timber, logging, and forestry workers
		498	Fishers, marine life cultivators, hunters, and kindred
Supervisors of mechanics and repairers	503	503	Supervisors of mechanics and repairers
Automobile mechanics and repairers	505	505	Automobile mechanics and repairers
		507	Bus, truck, and stationary engine mechanics
		508	Aircraft mechanics
		509	Small engine repairers
		514	Auto body repairers
Heavy equipement mechanics	516	516	Heavy equipement and farm equipment mechanics
		518	Industrial machinery repairers
		519	Machinery maintenance occupations
Repairers of electrical equipment	523	523	Repairers of industrial electrical equipment
		525	Repairers of data processing equipment
		526	Repairers of household appliances and power tools
		527	Telecom and line installers and repairers
		533	Repairers of electrical equipment, n.e.c.
		534	Heating, air conditioning, and refrigeration mechanics
Mechanics and repairers, n.e.c.	549	535	Precision makers, repairers, and smiths
		536	Locksmiths and safe repairers
		539	Repairers of mechanical controls and valves
		543	Elevator installers and repairers
		544	Millwrights
		549	Mechanics and repairers, n.e.c.
Supervisors of construction work	558	558	Supervisors of construction work
Carpenters and drywall installers	567	567	Carpenters
		573	Drywall installers
Electricians	575	575	Electricians
		577	Electric power installers and repairers
Plumbers, pipe fitters, and steamfitters	585	585	Plumbers, pipe fitters, and steamfitters
Misc. construction and related occupations	599	563	Masons, tilers, and carpet installers
		579	Painters, construction and maintenance
		583	Paperhangers
		584	Plasterers
		588	Concrete and cement workers
		589	Glaziers
		593	Insulation workers
		594	Paving, surfacing, and tamping equipment operators
		595	Roofers and slaters
		597	Structural metal workers
		598	Drillers of earth
		599	Misc. construction and related occupations
Extractive occupations	614	614	Drillers of oil wells
		615	Explosives workers
		616	Miners
		617	Other mining occupations
Production supervisors or foremen	628	628	Production supervisors or foremen
Metal and plastic workers	653	634	Tool and die makers and die setters
		637	Machinists
		643	Boilermakers
		649	Engravers
		653	Other metal and plastic workers
Dental laboratory and medical applicance technicians	678	678	Dental laboratory and medical applicance technicians

Precision and craft workers	684	644	Precision grinders and fitters
		657	Cabinetmakers and bench carpeters
		658	Furniture/wood finishers, other prec. wood workers
		666	Dressmakers, seamstresses, and tailors
		668	Upholsterers
		675	Hand molders and shapers, except jewelers
		677	Optical goods workers
		679	Bookbinders
		684	Other precision and craft workers
Butchers, bakers and batch food makers	686	686	Butchers and meat cutters
		687	Bakers
		688	Batch food makers
Plant and system operators	699	694	Water and sewage treatment plant operators
		695	Power plant operators
		696	Plant and system operators, stationary engineers
		699	Other plant and system operators
Woodworking and metal machine operators	703	645	Patternmakers and model makers
		703	Lathe, milling, and turning machine operatives
		706	Punching and stamping press operatives
		707	Rollers, roll hands, and finishers of metal
		708	Drilling and boring machine operators
		709	Grinding, abrading, buffing, and polishing workers
		713	Forge and hammer operators
		719	Molders and casting machine operators
		723	Metal platers
		724	Heat treating equipment operators
		727	Sawing machine operators and sawyers
		729	Nail, tacking, shaping and joining mach ops (wood)
		733	Other woodworking machine operators
Printing machine operators	736	734	Printing machine operators, n.e.c.
		736	Typesetters and compositors
Machine operators, n.e.c.	779	669	Nail, tacking, shaping and joining mach ops (wood)
		738	Winding and twisting textile and apparel operatives
		739	Knitters, loopers, and toppers textile operatives
		743	Textile cutting and dyeing machine operators
		744	Textile sewing machine operators
		745	Shoemaking machine operators
		747	Clothing pressing machine operators
		749	Miscellanious textile machine operators
		753	Cementing and gluing machne operators
		754	Packers, fillers, and wrappers
		755	Extruding and forming machine operators
		756	Mixing and blending machine operators
		757	Separating, filtering, and clarifying machine operators
		763	Food roasting and baking machine operators
		764	Washing, cleaning, and pickling machine operators
		765	Paper folding machine operators

		766	Furnance, kiln, and oven operators, apart from food
		769	Slicing, cutting, crushing and grinding machine
		774	Photographic process workers
		779	Machine operators, n.e.c.
		789	Painting and decoration occupations
Welders, solderers, and metal cutters	783	783	Welders, solderers, and metal cutters
Assemblers of electrical equipment	785	785	Assemblers of electrical equipment
Production checkers, graders, and sorters in	799	799	Production checkers, graders, and sorters in
Supervisors of motor vehicle transportation	803	803	Supervisors of motor vehicle transportation
Miscellanious transportation occupations	834	804	Truck, delivery, and tractor drivers
		808	Bus drivers
		809	Taxi cab drivers and chauffeurs
		813	Parking lot attendants
		823	Railroad conductors and yardmasters
		824	Locomotive operators: engineers and firemen
		825	Railroad brake, coupler, and switch operators
		829	Ship crews and marine engineers
		834	Miscellanious transportation occupations
Operating engineers of construction equipment	844	844	Operating engineers of construction equipment
Construction laborers	869	848	Crane, derrick, winch, hoist, longshore operators
		853	Excavating and loading machine operators
		859	Stevedores and misc. material moving occupations
		865	Helpers, constructions
		866	Helpers, surveyors
		869	Construction laborers
Laborers, freight, stock, and material handlers, n.e.c.	889	873	Production helpers
		875	Garbage and recyclable material collectors
		878	Machine feeders and offbearers
		885	Garage and service station related occupations
		887	Vehicle washers and equipment cleaners
		888	Packers and packagers by hand
		889	Laborers, freight, stock, and material handlers, n.e.c.