The Dual Local Markets: Family, Jobs, and the Spatial Distribution of Skills*

Jingting Fan† Ben Zou‡

This version: November 2021

Abstract

We study how the interactions between the local labor market and the local marriage market determine the spatial distribution of economic activities. We develop the first spatial equilibrium model with endogenous marriage formation. Calibrating the model to U.S. cities, we find that despite strong positive assortative matching, marriage is a force of spatial dispersion. Endogenous, heterogeneous returns from local marriage markets are the key driver of this result. Through the lens of the model, the secular decline in the preference for marriage, the increase in labor force participation among women, and the narrowing gender pay gap together explain between a third and a half of the spatial divergence in the U.S. between 1960 and 2000.

JEL codes: R12, R13, J12

Keywords: quantitative spatial models, spatial distribution of economic activities, local marriage market

*First draft, September 2020. For helpful comments, we thank Kerem Cosar, Jeremy Greenwood, Elisa Giannone, Chang-Tai Hsieh, Wen-Tai Hsu, Fernando Parro, Diana Van Patten, Ewout Verriest, and participants at the 2020 SMU Urban Economics Workshop, and the 2021 European Urban Economic Association Meetings. Xinle Pang provided excellent research assistance.

†Pennsylvania State University; jxf524@psu.edu

‡Michigan State University; benzou@msu.edu

1 Introduction

Despite the advancement in communication technologies, both jobs and family remain largely 'local'—most workers live with their spouses and still work within commuting distances from their homes. As such, the job and marital opportunities of young people are largely determined by what their cities have to offer. In turn, in choosing the city, people take into account their prospects in both local marriage and labor markets.

Through individuals' location choices, the two markets interact to shape the size and productivity of cities. Consider a city with plenty of skilled jobs. High-skill people are attracted to the city by job opportunities. As high-skill partners tend to be more sought-after in the marriage market, the resulting increase in the city's skill share makes it more attractive. Consequently, more people move to the city, affecting the equilibrium returns from the local marriage and labor markets. These mechanisms further interact with local spillovers, generating ripple effects.

The above scenario leads to three related questions. First, given the possible role of marriage prospects in location choice, does the marriage incentive make the spatial distribution of economic activities more concentrated? Second, how important are local spillovers and general equilibrium mechanisms? Third, the share of American adults who are married declined steadily across age groups since the 1960s; this period also saw an increasing concentration of population in the most skill intensive cities. Are these phenomena related? Answering these questions requires an equilibrium model with endogenous local marriage and labor markets. Although a recent literature (e.g. Redding, 2016; Caliendo et al., 2014) has developed quantitative models to examine various forces driving the spatial distribution of economic activities, the economic agents in these models are permanent singles, whose utility depends on local amenities and their own real wage. Differences in marriage market conditions across cities are not modeled.

In this paper we make two contributions. First, we develop a new quantitative spatial equilibrium model with endogenous marriage formation. We show that existing spatial models overlook the heterogeneity across cities in the 'appeal' of their local marriage markets, captured by a sufficient statistic, which could be an important force shaping not only the sorting of individuals across cities but also the response of the economy to changes in technology and policy. Second, we combine the model with the U.S. MSA-level data to answer the questions posed earlier. We find that although marriages are assortative—in the sense that people tend to have spouses with similar education levels—they are a force of spatial *dispersion*, i.e., without marriage, population and skills would have been more concentrated in already skill intensive cities. Moreover, the changing marriage institution between 1960 and 2000 can account for between a third and a half of the observed divergence among U.S. cities during this period.

To motivate the model, in Section 2, we provide descriptive evidence on the heterogeneous marriage patterns among U.S. cities and on the importance of such heterogeneity for location choices. In particular, we show that an individual's chance of marrying a high-skill spouse increases with the skill share of a city, which suggests that marriage outcomes likely depend on location choices. Moreover, although some cities, due to their occupation and industry structure, offer higher wages to females than other cities, the population gender ratios are similar across cities. This fact is consistent with individuals weighing the probability of matching with a spouse of the opposite gender in deciding where to move, but is at odds with the notion that location choices are made solely to maximize the real wage.

Section 3 describes the model. We embed a transferable utility marriage model in the spirit of Becker (1973, 1974) into a quantitative spatial equilibrium model in the tradition of the Rosen-Roback framework (Fajgelbaum and Gaubert, 2020). In the model, people differ by skill and gender. Young, single individuals consider where to start their career and family. They can move to any city after paying a migration cost. Once settled, individuals seek to match with a partner in the same city. To operationalize Becker's idea quantitatively, we follow the approach pioneered by Choo and Siow (2006) and assume that the utility for a couple consists of an *idiosyncratic* component that varies by person and a *systematic* component, which, as described below, summarizes the non-economic as well as the economic returns of a certain type of marriages. Given these returns, individuals determine the marriage outcome that gives them the highest utility. They can form a match with a potential partner by making endogenous transfers, which will act as prices to ensure the marriage market in each city is cleared. In equilibrium, a married couple divides up the marriage surplus between them via transfers, whereas those who remain single derive utility from their own consumption and local amenities.

The *systematic* marriage return encompasses several forces. The first, a non-economic component, is a 'taste' shifter specific to each type of marriages (i.e., whether a spouse is college educated or not). It reflects the broad social norm and is exogenous to the model. The second, the economic return, captures the impacts on household utility of income, amenities, and prices, all of which are equilibrium outcomes and vary by location. The married differ from the singles in this economic return in three ways: their income, their appreciation of local amenities which could differ by skill, and the option of having a stay-at-home partner producing home goods. The decision to stay at home, in turn, depends on the market price of home goods, spouse wages, and the 'taste' for home production, with the last reflecting the social norm. Given the economic components, disciplined by the data, we choose the average values of the two taste shifters to match the composition of marriages and the labor force participation rate of married women.[1]

In Section 4, we parameterize the model to match the U.S. spatial economy in 2000. We target the data in the *overall* composition of marriages and *city-specific* population, skill shares, and rents. The model successfully accounts for the vast heterogeneity in marriage market outcomes across cities. In particular, it generates empirically consistent cross-city patterns in the composition of marriages and the relative skill intensity between genders, despite none of these are targets of the calibration. This validates that the model captures the essence of the two local markets.

In Section 5, we examine the partial and general equilibrium impacts of marriage on the spatial distribution of economic activities. We show that the expected utility of a city for a given

[1] In the overwhelming majority of single-worker couples (98.7% in 1960, 85% in 2000), the husband works in the labor market and the wife works at home. Thus, in taking the model to the data, we assume for simplicity that the wife chooses whether to participate in the labor force while the husband always participates.

group of people (e.g. high-skill men) depends only on the utility of singles and the single rate among that group in the city, as below

$$\text{Expected utility} \propto \underbrace{\text{Utility from being single}}_{\text{local amenities+ real wage}} - \frac{\log(\text{single rate})}{\text{marriage pref. elasticity}}. \tag{1}$$

The second component is thus the sufficient statistic for the attractiveness of a city from the marriage incentive, which we call the marriage market premium. Intuitively, given the utility from being single, a higher single rate means being married in that city is relatively unattractive, so the marriage incentive is less important a reason for people to choose that city.

Equation (1) clarifies one departure of our model from workhorse quantitative spatial equilibrium models, in which the marriage market premium does not show up in the expected utility. In disciplining these models, researchers often back out amenities—which are not directly observed—as residuals to match the population distribution. This means that the measured amenities of a city would capture its marriage market premium. One might think this is only a matter of how amenities, a catch-all phrase, is defined. Yet, note that being an endogenous outcome that depends on the sorting of individuals, the marriage market premium might respond to shocks differently from other components of amenities. Thus, our model would imply different responses to shocks from existing models.

We examine the *partial equilibrium* impact of marriage on location choices by evaluating migration decisions using only the first component of the expected utility, effectively eliminating the marriage market premium for all cities.[2] Because in the data, as well as in the model, skill intensive cities have higher single rates among both high- and low-skill people, the elimination of the marriage market premium makes these cities more attractive for everyone. Because high-skill people face lower migration costs than low-skill people, however, more high-skill than low-skill people relocate to these cities. Consequently, the elimination of marriage market premia makes skill-intensive cities larger and even more skill intensive.

This result might at first appear surprising given the observation that in skill-intensive cities, high-skill people have a higher chance of being matched with a high-skill spouse (Compton and Pollak, 2007). If matching with a high-skill spouse is preferable, the presence of marriage incentive would seem to attract high-skill people to skill intensive cities. Why, then, would the elimination of marriage lead to more concentration? We show that although marrying a high-skill spouse gives a high-skill person more utility than being single, such advantages are smaller in skill intensive cities due to a smaller marriage surplus in these cities. In addition, the relative scarcity of low-skill partners in these cities means that the high-skill people ending up marrying a low-skill spouse will not be able to extract as favorable a division of the marriage surplus as in other cities. So even though in skill intensive cities the chance of forming a 'power couple'

[2]This thought experiment is equivalent to setting the single rate to one in all cities. This is a partial equilibrium exercise because by keeping the utility from being single fixed at the baseline equilibrium values, it overlooks the endogenous responses in local labor and marriage markets and through local spillovers.

is higher, the marriage market premium is not necessarily higher. This result underscores the importance modeling marriage market returns as endogenous outcomes.

To understand the *general equilibrium* impact of marriage, we shut down marriages entirely by setting the non-economic marriage surplus term to a sufficiently negative number. As before, we find that skill intensive cities grow in both size and skill intensity. In particular, population in cities like Washington, D.C. and San Francisco grow by more than 50%. The gradient of the change in log population with respect to the baseline log skill intensity (the ratio between high- and low-skill people counts) is 1.55, and the gradient of the change in log skill intensity with respect to the baseline log skill intensity is 1.75. Both differ substantially from the results of the partial equilibrium exercise (0.37 and 2.2, respectively), highlighting the importance of the general equilibrium effects.

In Section 6, we examine the implications of the changing marriage institution between 1960 and 2000 for the U.S. spatial economy. Over this period, the share of people between ages 25 and 54 who are married decreased steadily from 83% to 68%; around the same time, the U.S. economy also experienced increasing spatial divergence, with skill intensive cities outgrowing others in size and in skill intensity.

We calibrate the model to the 1960 U.S. economy, in doing so we decompose the differences between 1960 and 2000 into changes in the 'tastes' for marriage and home production, and other fundamentals. Among other things, we find that over this period, there has been a decrease in the 'taste' for home production, and a decrease in the 'taste' for all marriages except marriages between a low-skill man and a high-skill woman. Both reflect changes in social norms.

We feed these changes into the 1960 economy to evaluate their impacts. We find that the decrease in the 'taste' for marriage is the most important contributor to the declining share of married people. With marriage becoming less attractive, fewer people are now bonded by marriage, and more move to skilled cities. This process is amplified by local spillovers. On the other hand, the decrease in the value of home production, which leads to an increase in female labor force participation, has only a small direct effect. When we feed into the model changes in all fundamental parameters, the most skill intensive cities in 1960 see around 15% population growth and a 10 p.p. increase in skill intensity. The overall impact of these changes, summarized by the gradients of log population growth and the change in city skill intensity with respect to the initial log skill intensity of cities, is 0.20 for population and 0.054 for skill intensity. These two gradients are half and one-third of their empirical counterparts, respectively. Thus, the evolving marriage institution promises to be an important driving force of the observed changes in within-country economic geography.[3]

This paper is related to three strands of the literature. First, we build on a rapidly growing literature on quantitative spatial models, reviewed by Redding and Rossi-Hansberg (2017). The spatial equilibrium block of our model and our outcome of interest—the distribution of economic

[3]While this paper focuses on the United States, Appendix Table A.1 shows that the declining marriage rate went hand in hand with increasing population concentration in most industrialized countries in the past few decades.

activities across space—are shared by a number of studies focusing on a range of different forces such as sorting by skill and occupation (Davis and Dingel, 2019; Fajgelbaum and Gaubert, 2020; Rossi-Hansberg, Sarte and Schwartzman, 2019), domestic transport infrastructure (Donaldson, 2018; Fajgelbaum and Schaal, 2019; Allen and Arkolakis, 2019), productivity spillovers and firm sorting (Gaubert, 2018), and commuting decisions (Monte, Redding and Rossi-Hansberg, 2018).[4] Our contribution is to develop a tractable model with endogenous local marriage markets, which we show are first-order determinants of the size and skill composition of cities. This model is suitable for analyzing the spatial impacts of policies and shocks that are either gender non-neutral (e.g., the transition to the service economy), or might have differential impacts on married and single households (e.g., changes in the tax code or regulations affecting the relative price of owning or renting homes).

Second, in recent decades, the U.S. economy has experienced a reshuffling of economic activities across space, resulting in salient trends including spatial divergence, suburbanization, and gentrification of city centers. In explaining these trends, the literature has proposed explanations based on housing constraints (Hsieh and Moretti, 2019), reductions in commuting costs (Redding, 2021), endogenous amenities (Diamond, 2016), rising income inequality (Couture, Gaubert, Handbury and Hurst, 2019), and changing production and trade technologies (Giannone, 2017, Eckert, 2019, Jiao and Tian, 2019). Enabled by our new model, we investigate a different explanation: the declining marriage institution. Our focus is related to Costa and Kahn (2000), who argue that the increase in the share of dual-career couples is responsible for the concentration of 'power couples' in large cities. Disciplined by the data, our model suggests that although this channel is present, its impact is modest; the decline in the marriage rate is quantitatively much more important.

Finally, our model builds on the quantitative transferable utility marriage framework (Choo and Siow, 2006). The studies of equilibrium marriage models have evolved around two strands: first, a micro economic strand primarily interested in the identification and estimation of marital preferences and their determinants (Graham, 2013; Chiappori, Salanié and Weiss, 2017); second and more closely related, a macro strand on the interplay between marriage and aggregate outcomes, such as female labor force participation and fertility (see Greenwood, Guner and Vandenbroucke, 2017 for a survey). Both strands generally focus on geographically aggregated marriage markets, despite that several reduced-form studies have shown that marriage outcomes differ significantly across space and argued this to be important for location choices.[5] Our contribution is to extend a workhorse matching model into a multi-region setting and to demonstrate that it

[4] Also related, existing research has used the same class of models to study the transmission of productivity shocks across space (Caliendo, Parro, Rossi-Hansberg and Sarte, 2014) and to quantify the impact of trade and immigration shocks on the location choice and welfare of people (Artuç, Chaudhuri and McLaren, 2010; Caliendo, Dvorkin and Parro, 2019; Burstein, Hanson, Tian and Vogel, 2020).

[5] For example, Costa and Kahn (2000) and Compton and Pollak (2007) show that power couples are increasingly concentrated in large cities and propose explanations for this phenomenon; Edlund (2005) hypothesizes that large cities attract single women (especially the low-skilled) because these cities offer greater returns in the marriage market.

goes a long way towards explaining the spatial heterogeneity in marriage outcomes.[6]

2 Spatial Heterogeneity in Local Marriage Markets

U.S. cities differ markedly from one another in the outcomes of their local marriage markets. In the Washington, D.C. metropolitan area, for example, about 37% of married couples in 2000 are 'power couples'—couples in which both the husband and the wife are college graduates; in El Paso, TX, the share of such couples is only 10%. Such heterogeneity in marriage outcomes is inherently related to the spatial distribution of skills. Figure 1a plots the probability among individuals between ages 25 and 54 of marrying a high-skill partner against the log skill ratio in the local market, separately for men and women with and without a college degree. The fitted lines show a clear pattern: for all types of individuals, the probability of marrying a high-skill spouse increases with the skill ratio of the local market. The increase can be substantial. For a high-skill man, for example, living in a city with twice as many (in ratio) high-skill people is associated with a 10 percentage point (p.p.) higher probability of marrying a high-skill woman.[7]

Given the large difference in marriage market outcomes, it is only natural that single young people take into account their marriage prospects in deciding where to live. Figure 1b provides suggestive evidence for this hypothesis. The figure plots the log gender ratios of cities against the wage differences between men and women. It shows that even though some cities offer a higher wage to men relative to women than others, these cities do not have a higher share of men.

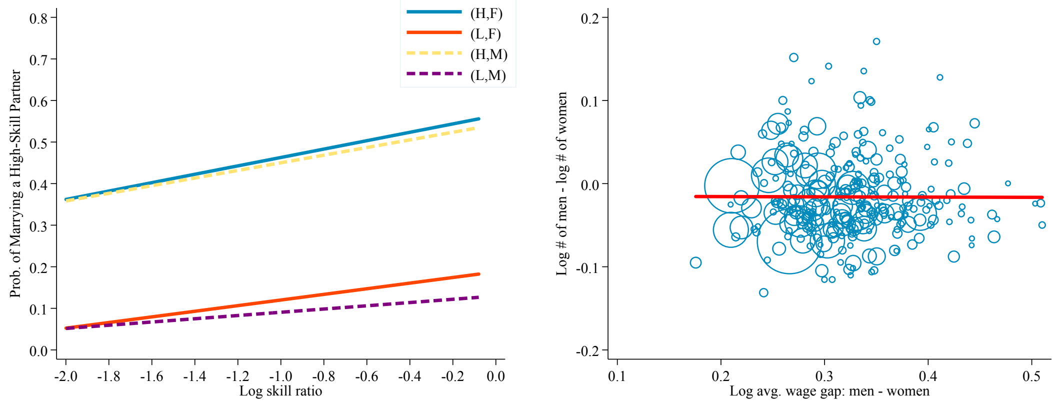
Such a pattern is at odds with workhorse spatial equilibrium models, in which individuals make location choices solely to maximize the (amenity-adjusted) real returns from the labor market, and which would predict that relative labor supply increases in relative wages.[8,9] On the other hand, it is supportive of marriage prospects being an important factor in location

[6] With the focus on the spatial and general equilibrium implications of the migration decision of young adults, our paper has less in common with existing partial equilibrium studies on how the 'two-body problem' affects the migration decision of dual-career households (Gemici, 2011; Venator, 2021). While it is conceptually straightforward to extend the model to incorporate a second-stage migration for married people, our focus on the choice of young adults is guided by the questions we seek to answer. Ultimately, our outcome of interest is the distribution of economic activities across space; in the data, young, single adults—much more mobile than older, married couples—account for the majority of inter-city migration. For example, according to the 2000 Census, 61% of college-educated singles between ages 25 and 29 migrate between MSAs in the past 5 years; among those who are married between ages 45 and 49, this share is merely 15%.

[7] Of course, this cross-sectional pattern could be due to power couples moving disproportionally to cities with a high skill share. Using panel data, Compton and Pollak (2007) conclude that this alternative hypothetical does not have strong empirical support and that the observed patterns are better explained by higher rates of power couple formation in these cities.

[8] A concern with this interpretation of the evidence is that gender wage differences are endogenous outcomes and a function of relative supply, i.e., Figure 1b can be tracing out a demand curve instead of a supply curve. To address this concern, Appendix Figure A.1a shows a version of this graph where gender wage differences are predicted based solely on the industry and occupation composition of the local market. It shows a similar pattern: local labor markets in which men have a wage premium due to the demand side do *not* have more men than women. This pattern is also not confounded by the city size or city skill ratio, which might independently affect the location choice of men versus women. In fact, gender wage differences are not systematically associated with either size or skill intensity. Appendix Figure A.1b shows that adjusting gender wage differences for city size and skill ratio results in a similar pattern.

[9] In principle, one can specify city and gender-skill specific amenities to match any kind of population distribution; in fact, one can specify amenities to be specific to each city-by-marriage type to match the spatial heterogeneity in marriage outcomes. We do *not* follow this path. Instead, in quantification we will let the marriage market forces determine these patterns.



(a) Prob. of Marrying a Skilled Spouse

(b) Gender Ratio and Relative Wage

Figure 1: Local Marriage Market and Skill Distribution

Note: Each unit is an MSA in 2000. The sample includes individuals between ages 25 and 54. Panel A shows the fitted linear lines between the probability of marrying a high-skill spouse and the log skill ratio in the local market. The skill ratio is defined as the number of high-skill workers over the number of low-skill workers. H and L indicate skill level; F and M indicate gender. Panel B shows the scatter plot and the fitted linear line between the local log gender ratio and the local relative wage. The relative wage is calculated using the annual earnings of full-time workers, accounting for education and potential experiences through Mincerian adjustments. The slope for the fitted line is 0.03, with a robust standard error of 0.05.

choices—people move to cities where their gender is relatively scarce in the marriage market, thus blurring the impact of relative wages on relative local labor supply. In the rest of this paper, we develop a spatial equilibrium model in which marriage markets are heterogeneous across space and individuals take such heterogeneity into account in their location choices.

3 Model

We combine the Rosen-Roback spatial equilibrium model and a Beckerian competitive matching model in a tractable framework for quantitative investigation. This section describes the model setup.

3.1 Environment

The model economy is static and consists of N cities, denoted by $n = \{1, 2, ..., N\}$. It is populated by four types of residents: males and females with high and low skill levels. We denote the type of a person by (s, e), where $s \in \{F, M\}$ stands for gender and $e \in \{H, L\}$ stands for education. Each city is endowed with an exogenous number of each type; within a type, everyone is ex-ante identical. Cities differ in productivity and amenities, both of which depend on an exogenous fundamental component and an endogenous component that is shaped by agglomeration forces.

Residents in the economy start as single young adults, who maximize their expected utility by choosing the city to start their family and career. In the chosen city, single men and women meet to form families (or decide to remain single). Finally, given local wages and their preferences, households make consumption and home production decisions.

Our description of the model proceeds as follows. We start with the location choice of young adults and then describe the setting for and the outcome of local marriage markets, taking as given the household utility in each city. We then describe household decisions and the spatial model, which provides a micro-foundation for household utility and specifies how different forces interact in general equilibrium.

3.2 Location Choices of Young Adults and Local Marriage Markets

Location choices of young adults. At the beginning of time, young adults from city o, all single, choose a city to reside. Their location choice takes into account a few factors: the migration cost from the birthplace to the destination city, the expected utility of living in a city, which summarizes its marriage and labor market prospects, and an idiosyncratic preference draw for that city. Formally, a young adult from city o chooses d to maximize:

$$\max_d(\overline{U}^e_{d,s} - d^e_{od,s} + \tilde{\epsilon}^e_{d,s}),$$

in which $\overline{U}^e_{d,s}$ is the expected utility of a person in city d with gender s skill e, $d^e_{od,s}$ is the migration cost between cities o and d, and $\tilde{\epsilon}^e_{d,s}$ is an idisocynratic location-specific utility draw generated from a Gumble distribution with dispersion parameter θ^e_s.

Let the number of young adults of type (s,e) born in city o be $L^e_{o,s}$. The number of people moving from city o to city d is $L^e_{o,s} \cdot \pi^e_{od,s}$, with $\pi^e_{od,s}$ defined by

$$\pi^e_{od,s} = \frac{\exp\left(\theta^e_s(\overline{U}^e_{d,s} - d^e_{od,s})\right)}{\sum_d \exp\left(\theta^e_s(\overline{U}^e_{d,s} - d^e_{od,s})\right)}. \tag{2}$$

Local marriage markets. After settling into a city, young adults participate in the local marriage market. Based on the skill composition of the spouses, there are four types of marriages. We denote a marriage by (e, e'), with e being the skill of the husband and e' being that of the wife. Slightly abusing notation, we will use $(e, \varnothing)$ to denote a single man with skill e and $(\varnothing, e')$ to denote a single woman with skill e'.

In deciding whether and whom to marry, individuals consider their utility from a marriage and from staying single. Their utility from either outcome is the sum of two components. The first is systematic and specific to each type of outcome. It summarizes both the economic and non-economic returns to a marriage and will be determined by household consumption and labor supply decisions. We use $\overline{V}^{e,\varnothing}_d$ and $\overline{V}^{\varnothing,e'}_d$, $e, e' \in \{H, L\}$ to denote this deterministic component of utility for single men and women in city d, respectively; we use $\overline{V}^{e,e'}_d$, $e, e' \in \{H, L\}$ to denote the systematic component of the utility of type (e, e') households in city d.

The second component is idiosyncratic and captures the personal taste for being single and for marrying a spouse of certain skill. We think of these idiosyncratic draws as reflecting the preference for spouse personalities that might be correlated with spouse skill. Motivated by the observation that most people learn their marital preferences over time and through dating experiences, we assume that these draws realize after individuals settle into a city. Let $\omega \in \Omega_d^{M,e}$ denote a man in city d with skill e and let $\xi_M^{e,e'}(\omega)$, $e' \in \{H, L, \varnothing\}$ denote the realization of the taste draws of man ω for a marriage market outcome, e'. Similarly, let $\xi_F^{e,e'}(\omega')$ denote the idiosyncratic preference of a woman $\omega' \in \Omega_d^{F,e'}$ for the three possible marriage market outcomes: $e \in \{H, L, \varnothing\}$.

Factoring in both the deterministic and the idiosyncratic components, the *household* utility for a marriage between $\omega \in \Omega_d^{M,e}$ and $\omega' \in \Omega_d^{F,e'}$ is

$$\overline{V}_d^{e,e'} + \xi_M^{e,e'}(\omega) + \xi_F^{e,e'}(\omega').$$

Similarly, the utility of a man ω and a woman ω' from staying single is, respectively,

$$\overline{V}_d^{e,\varnothing} + \xi_M^{e,\varnothing}(\omega) \text{ and } \overline{V}_d^{\varnothing,e'} + \xi_F^{\varnothing,e'}(\omega').$$

Given their idiosyncratic preference draws, each individual finds the arrangement that maximizes their own utility. Following the literature on competitive matching in the marriage market, we make two assumptions. First, the idiosyncratic component of the utility is independent across two spouses, i.e., $\xi_M^{e,e'}(\omega)$ and $\xi_F^{e,e'}(\omega')$ are independent. By ruling out the interaction between the idiosyncratic taste of the husband and that of the wife, this assumption allows us to identify the non-economic component in $\overline{V}_d^{e,e'}$ using marriage outcomes (see Galichon and Salanié, 2020 for a discussion of identification in transferable utility models). Second, utility is transferable between the spouses. This assumption implies that the utility of a marriage, $\overline{V}_d^{e,e'} + \xi_M^{e,e'}(\omega) + \xi_F^{e,e'}(\omega')$, will be fully divided between the spouses.

The outcome of a local marriage market is a match, which prescribes who stays single, who marries whom, and how the marriage surplus is split between the husband and the wife.[10] A match is stable, if no one can be better off by deviating from it (e.g., by marrying a different person or by staying single). Formally, a stable match in city d is defined as below:

Definition 1. *A stable match in city d is a measure μ_d defined on $(\Omega_d^{M,e} \cup \varnothing) \times (\Omega_d^{F,e'} \cup \varnothing)$, $e, e' \in \{H, L\}$ and a set of payoffs for any man ω and any woman ω', denoted $u_d^{M,e}(\omega)$ and $u_d^{F,e'}(\omega')$, such that:*

1. *The marginal of μ_d over $\Omega_d^{M,e}$ and $\Omega_d^{F,e'}$ is the measure for distribution of all males and females in city d: $\mu_d^{M,e}$ and $\mu_d^{F,e'}$, respectively.*
2. $u_d^{M,e}(\omega) \geq \overline{V}_d^{e,\varnothing} + \xi_M^{e,\varnothing}(\omega), \forall \omega \in \Omega_d^{M,e},\ e \in \{H, L\}$
3. $u_d^{F,e'}(\omega') \geq \overline{V}_d^{\varnothing,e'} + \xi_F^{\varnothing,e'}(\omega), \forall \omega' \in \Omega_d^{F,e'},\ e' \in \{H, L\}$

[10] We do not specify the protocol for the marriage market. One way for the marriage market to operate is to allow each person to make a transfer to his/her partner as part of the proposal. Such transfers will then act as prices that clear the marriage market (see Becker, 1973; Choo and Siow, 2006). In some cultures, such transfers take the form of dowry. In other cultures where direct payments for marriages is not a social norm, such transfers could take the form of a commitment to future household decisions in a way that is more aligned with one spouse's preference. For example, to court a partner, one might promise to do more household chores in the future.

4. $u_d^{M,e}(\omega) + u_d^{F,e'}(\omega') = \overline{V}_d^{e,e'} + \xi_M^{e,e'}(\omega) + \xi_F^{e,e'}(\omega)$, $\forall(\omega,\omega')$ *in the support of* μ_d

5. $u_d^{M,e}(\omega) + u_d^{F,e'}(\omega') \geq \overline{V}_d^{e,e'} + \xi_M^{e,e'}(\omega) + \xi_F^{e,e'}(\omega')$, $\forall(\omega,\omega') \in \Omega_d^{M,e} \times \Omega_d^{F,e'}$.

The first condition states that the match should assign every man and woman either a partner or a single status, but does not assign roles to more men and women than available in the city. The stable match ensures everyone a level of utility. The second and third conditions state that the assigned utility should be weakly higher than that from staying single—i.e., not participating in the match at all—for everyone. It can also be viewed as a participation constraint. The fourth condition ensures that for marriages formed according to the assignment, the utility of the two spouses should add up to the household utility, an implication of the transferable utility assumption. Finally, the fifth condition says that for *any two* people of the opposite genders, the joint utility if they are matched should not be higher than the sum of their promised utility according to the assignment. This condition ensures that no two persons can increase their utility above the levels in the stable match by forming a new union.

We solve for the stable match for quantitative analysis.[11] In principle, a stable match is a high-dimensional object encompassing the outcomes of all individuals and their utility. Lemma 1 and Proposition 1 show that we can characterize it parsimoniously.

Lemma 1. *In a stable match, there exist* $U_{d,M}^{e,e'}$, $U_{d,M}^{e,\varnothing}$, $U_{d,F}^{e,e'}$, $U_{d,F}^{\varnothing,e'}$ *for all* $e \in \{H,L\}$ *and* $e' \in \{H,L\}$, *in total 12 scalers for each city* d, *such that the utility assigned to individuals is as follows:*

1. *For men and women staying single in the stable match, their assigned utility satisfies*

$$u_d^{M,e}(\omega) = U_{d,M}^{e,\varnothing} + \xi_M^{e,\varnothing}(\omega), \;\; u_d^{F,e'}(\omega') = U_{d,F}^{\varnothing,e'} + \xi_F^{\varnothing,e'}(\omega'). \tag{3}$$

2. *For men and women assigned to a marriage* (e,e') *in the stable match, their assigned utility satisfies*

$$\begin{aligned} u_d^{M,e}(\omega) &= U_{d,M}^{e,e'} + \xi_M^{e,e'}(\omega) \\ u_d^{F,e'}(\omega') &= U_{d,F}^{e,e'} + \xi_F^{e,e'}(\omega'). \end{aligned} \tag{4}$$

That is, the utility of an individual with gender s *in the stable match is the sum of the scaler* $U_{d,s}^{e,e'}$ *and his/her own taste draw for outcome* (e,e'); *it does not depend on the taste shock of his/her partner.*

3. *These 12 scalers split the systematic parts of household utility, i.e.,*

$$\begin{aligned} U_{d,M}^{e,\varnothing} &= \overline{V}_d^{e,\varnothing} \\ U_{d,F}^{\varnothing,e'} &= \overline{V}_d^{\varnothing,e'} \\ U_{d,M}^{e,e'} + U_{d,F}^{e,e'} &= \overline{V}_d^{e,e'}. \end{aligned} \tag{5}$$

[11] The literature has shown that in settings like ours, a stable match generally exits and is unique (Decker et al., 2013; Galichon and Salanié, 2020).

The proof of the lemma is delegated to Appendix B. Lemma 1 shows that, in a stable match, the utility of an individual is simply the sum of that individual's own idisocynratic preference for his/her outcome and a scaler independent of individual preferences. Moreover, the scalers of two spouses split the systematic component of the utility from a marriage. This result allows us to characterize the equilibrium outcome of each individual as a simple discrete choice problem, as described in the following proposition:

Proposition 1. *In a stable match, for a man $\omega \in \Omega_d^{M,e}$, his assigned utility $u_d^{M,e}(\omega)$ satisfies:*

$$u_d^{M,e}(\omega) = \max_{e' \in \{H,L,\varnothing\}} \{U_{d,M}^{e,e'} + \xi_M^{e,e'}(\omega)\}. \tag{6}$$

In addition, the outcome of ω in the stable match is the solution to the above maximization problem. Similarly, the assigned utility of a woman $\omega' \in \Omega_d^{F,e'}$ satisfies:

$$u_d^{F,e'}(\omega') = \max_{e \in \{H,L,\varnothing\}} \{U_{d,F}^{e,e'} + \xi_F^{e,e'}(\omega)\}, \tag{7}$$

and the outcome of ω' in the stable match solves the maximization problem in equation (7).

According to Proposition 1, although individuals can make arbitrary transfers to attract the desired type of partners, who also have their own preferences, the outcome of an individual boils down to a one-side discrete choice problem determined by a few scalers, $U_{d,M}^{e,e'}$, $U_{d,M}^{e,\varnothing}$, $U_{d,F}^{e,e'}$, and $U_{d,F}^{\varnothing,e'}$, which are equilibrium outcomes and seen as exogenous by individuals. In a sense, these systematic utility components act as prices that respond to the 'tightness' of the marriage market for different types of individuals, ensuring that no one has an incentive to deviate.[12]

For tractability, we follow the literature (e.g. Choo and Siow, 2006; Chiappori et al., 2017) and make the following parametric assumption:

Assumption 1. *The idiosyncratic preference of individual ω, denoted by $\vec{\xi}_M^e(\omega) \equiv \left(\xi_M^{e,H}(\omega), \xi_M^{e,L}(\omega), \xi_M^{e,\varnothing}(\omega)\right)$ is drawn i.i.d. from a Gumble distribution with parameter κ_M^e and a cumulative distribution function of $F(x) = \exp(-\exp(\kappa_M^e \cdot x))$; similarly, the idiosyncratic preference of a female with skill e' over the three marriage market outcomes is drawn i.i.d. from a Gumble distribution with parameter $\kappa_F^{e'}$.*

Under Assumption 1, equations (6) and (7) lead to the expressions for the fraction of men and women with a certain marriage market outcome as:

$$\begin{aligned} r_{d,M}^{e,e'} &= \frac{\exp(\kappa_M^e U_{d,M}^{e,e'})}{\sum_{e'' \in \{H,L,\varnothing\}} \exp(\kappa_M^e U_{d,M}^{e,e''})}, \ \forall e' \in \{H,L,\varnothing\}, \forall e \in \{H,L\} \\ r_{d,F}^{e,e'} &= \frac{\exp(\kappa_F^{e'} U_{d,F}^{e,e'})}{\sum_{e'' \in \{H,L,\varnothing\}} \exp(\kappa_F^{e'} U_{d,F}^{e'',e'})}, \ \forall e \in \{H,L,\varnothing\}, \forall e' \in \{H,L\}. \end{aligned} \tag{8}$$

[12] To see how these scalers can be interpreted as prices, assume the matching protocol is for men to propose. If in equilibrium a man ω matches with a woman ω', then he would optimally propose to give $U_{d,F}^{e,e'} + \xi_F^{e,e'}(\omega')$ to her while keep the remaining household utility $[V_d^{e,e'} + \xi_F^{e,e'}(\omega') + \xi_M^{e,e'}(\omega)] - [U_{d,F}^{e,e'} + \xi_F^{e,e'}(\omega')] = U_{d,M}^{e,e'} + \xi_M^{e,e'}(\omega)$ to himself.

The first line of the equation is the fractions of type e men in three outcomes: marrying a high-skill woman, marrying a low-skill woman, or remaining single. The second line is the corresponding expressions for women with skill e'.

Equation (8) relates marriage outcomes to systematic components of utilities and the dispersion parameter of idiosyncratic draws, with the latter governing how sensitive individual choices are to systematic utilities. For instance, a large value for κ_F^L means the idiosyncratic component weighs little in the marriage decision of low-skill women, so the marriage market outcomes of low-skill women should respond strongly to the systematic returns. A small value for κ_F^L instead means that an individual's observed type—her skill—does not have much explanatory power over their preferences. Therefore, the differences in the systematic returns would not lead to large changes in their marriage market outcomes.

Equation (8) also highlights an identification challenge: in a setting with *one* marriage market, the parameters governing the Gumble distributions and the utility in each type of marriage enter the expression in product, and thus cannot be separately identified from marriage market outcomes alone. This has led earlier studies in this literature, such as Choo and Siow (2006), to normalize the Gumble parameters. By embedding the marriage decision in a spatial equilibrium model, we will be able to exploit the restriction imposed by location choices—in combination with the variation in marriage market outcomes across cities—to identify all Gumble parameters. We will return to this in the next section.

In equilibrium, individual choices need to be consistent with market outcomes. Let $\tilde{N}_{d,s}^e$, for $s \in \{M, F\}$, $e \in \{H, L\}$ be the number of (s, e) type of people choosing to reside in d, we have:

$$\tilde{N}_{d,M}^{e} \cdot r_{d,M}^{e,e'} = \tilde{N}_{d,F}^{e'} \cdot r_{d,F}^{e,e'}, \ \forall e, e' \in \{H, L\}. \tag{9}$$

Equation (9) can be viewed as the marriage market clearing condition in a decentralized competitive equilibrium. Thus, given κ_s^e, $\tilde{N}_{d,s}^e$, and the utility associated with different types of single and married households ($\overline{V}_d^{e,e'}$, $\overline{V}_d^{e,\varnothing}$, $\overline{V}_d^{\varnothing,e'}$), a stable match of the marriage market in city d is characterized by $U_{d,M}^{e,e'}$, $U_{d,M}^{e,\varnothing}$, $U_{d,F}^{e,e'}$, $U_{d,F}^{\varnothing,e'}$, such that equations (5) and (9) are satisfied (in total 12 equations and 12 unknowns for each city).

We use the model to examine how the marriage incentive affects the location choice of young adults. Note that under Assumption 1, Proposition 1 implies that the expected utility of type e men and type e' women from living in city d, respectively, is

$$\begin{aligned} \overline{U}_{d,M}^{e} &= \frac{\bar{\gamma}}{\kappa_M^e} + \frac{1}{\kappa_M^e} \log\Big(\sum_{e' \in \{H,L,\varnothing\}} \exp(\kappa_M^e U_{d,M}^{e,e'}) \Big) \\ \overline{U}_{d,F}^{e'} &= \frac{\bar{\gamma}}{\kappa_F^{e'}} + \frac{1}{\kappa_F^{e'}} \log\Big(\sum_{e \in \{H,L,\varnothing\}} \exp(\kappa_F^{e'} U_{d,F}^{e,e'}) \Big), \end{aligned} \tag{10}$$

in which $\bar{\gamma}$ is the Euler-Mascheroni constant. Combing this with equations (5) and (8) gives us:

$$\overline{U}^{e}_{d,M} = \frac{\bar{\gamma}}{\kappa^{e}_{M}} + \overline{V}^{e,\varnothing}_{d} - \frac{1}{\kappa^{e}_{M}} \log(r^{e,\varnothing}_{d,M}) \tag{11}$$
$$\overline{U}^{e'}_{d,F} = \frac{\bar{\gamma}}{\kappa^{e'}_{F}} + \overline{V}^{\varnothing,e'}_{d} - \frac{1}{\kappa^{e'}_{F}} \log(r^{\varnothing,e'}_{d,F}),$$

where $r^{e,\varnothing}_{d,M}$ and $r^{\varnothing,e'}_{d,F}$ are, respectively, the single rate among (M, e) and (F, e') in city d, and $\overline{V}^{e,\varnothing}_{d}$ and $\overline{V}^{\varnothing,e'}_{d}$ are the utility from staying single for (M, e) and (F, e').

The marriage market premium. Equation (11) shows that compared with the workhorse quantitative spatial model (e.g., Redding, 2016), in which everyone is assumed to be single and makes decisions based on their own utility, in our model, an extra term depending on the single rate enters the expected utility. This difference has two key implications.

The first implication is on the theory-based measurement of amenities. Since the workhorse model omits the variation in the probability of being married, when calibrated to match the same population distribution to back out amenities, it underestimates the amenities of cities with a high single rate and overestimates the amenities of cities with a low single rate.[13] To researchers who are not primarily concerned about the exact interpretation of measured amenities, this difference might seem a matter of labeling. Note, however, that single rates are equilibrium outcomes that depend on the sorting of people and will respond to changes in the local economy differently from other components of amenities (such as the supply of theaters and restaurants). Our model therefore implies different responses to changes in policy or technology, even if these changes are not directly related to marriage.

Second, according to equation (11), conditional on the utility from being single, the probability of an individual ending up single captures the attractiveness of the city's marriage market. Good marriage market prospects imply a lower single rate, hence a higher overall expected utility. The importance of preference heterogeneity, summarized by κ^{e}_{s}, modulates variation in utility across cities due to variation in single rates.

Thus, $\frac{1}{\kappa^{e}_{M}} \log(r^{e,\varnothing}_{d,M})$ and $\frac{1}{\kappa^{e'}_{F}} \log(r^{\varnothing,e'}_{d,F})$ are the sufficient statistics for the *partial equilibrium* marriage market premium of a city—i.e., how much a city's utility stems from it being a good place for people to meet their ideal spouses. By setting the single rate of all cities in equation (11) to one and plugging the counterfactual expected utility into equation (2), we can examine how location choices change in response to the elimination of marriage incentives. With $\overline{V}^{e,\varnothing}_{d}$ and $\overline{V}^{\varnothing,e'}_{d}$ fixed, such a thought experiment is valid only in partial equilibrium. Indeed, with a change in the marriage incentive, general equilibrium interactions among labor, housing, the marriage markets will lead to changes in $\overline{V}^{e,\varnothing}_{d}$ and $\overline{V}^{\varnothing,e'}_{d}$. The rest of this section describes household decisions and the spatial equilibrium model incorporating exactly these mechanisms.

[13] A related but different approach that recovers the importance of various local factors in location choices is to regress housing prices on measures of city amenities (Rosen, 1974; Nelson, 1978). Under the assumption that amenities are priced into housing, such regressions recover the valuation of individuals for these amenities. Equation (11) suggests that such regressions can also be biased due to the omitted variable.

3.3 Consumption and Female Labor Force Participation

We start with consumption and labor market decisions, made after the marriage market has concluded. In the data, people who are past the marriage stage tend to have a lower migration rate than younger people. For simplicity, we assume people work in the city they choose as young adults.[14] We describe first the decisions of single households and then those of couples.

Decisions of singles. Single households in city d enjoy an indirect utility of $\overline{V}^{e}_{d,s}$, which is obtained as a solution to the problem below:

$$\overline{V}^{e}_{d,s} \equiv \max_{h,n} \log\left(A^{e}_{d} \cdot (I^{e}_{d,s} - r_d \cdot h - p_n \cdot n)^{(1-\alpha-\beta)} \cdot h^{\alpha} \cdot n^{\beta}\right).$$

Inside the log utility function, A^{e}_{d} is the amenities of city d to a person with skill e; h is the quantity of housing consumed and α is the housing share in consumption; n is the consumption of home goods and β is its share. As will be explained below, for married couples, these home goods could be produced by a stay-at-home partner, but for singles, the only way to obtain such goods is buying from the market at price p_n. Examples of market alternatives to home goods include food away from home, meal and grocery deliveries, and laundry and house cleaning services. $I^{e}_{d,s}$ is the income of an individual with skill e and gender s, which comes from two sources: labor earnings $W^{e}_{d,s}$ and a lump-sum transfer t from the government, i.e.,

$$I^{e}_{d,s} = W^{e}_{d,s} + t.$$

$(I^{e}_{d,s} - r_d \cdot h - p_n \cdot n)$ is therefore the consumption of regular goods after paying for housing and home goods. Under the optimal choice of n and h, the indirect utility is given by:

$$\overline{V}^{e}_{d,s} = c_1 + \log(A^{e}_{d}) + \log(I^{e}_{d,s}) - \alpha \log(r_d) - \beta \log(p_n), \tag{13}$$

where $c_1 = \log\left((1-\alpha-\beta)^{1-\alpha-\beta}\alpha^{\alpha}\beta^{\beta}\right)$ is a constant.

Decisions of married couples. In city d, couples decide whether the wife would participate in the labor market.[15] Specifically, the wife receives an idiosyncratic taste shock for each of the two choices, to work or to stay at home, and chooses the one to maximize the household utility. We denote these shocks ζ^{H} for home production and ζ^{W} for work. For wives with skill level e', their draws will be from the Gumble distribution with dispersion parameter $\eta^{e'}_{F}$. Concretely,

[14] For local marriage and labor markets to have meaningful interactions, the main assumption we need is that after marriage, people cannot move frictionlessly across MSAs. An earlier version of this paper incorporates a second-stage migration decision for after the marriage market. We find that when calibrated to match the observed migration rate, which is far from the case of frictionless mobility, that model generates similar predictions.

[15] Due to the norm from the patriarchal history of the human society, home production by stay-at-home wives has been much more common than that by stay-at-home husbands. To capture the empirically more important margin and for simplicity, we allow wives, but not husbands, to make a labor force participation decision. We model such decisions for married but not single women because the change in labor force participation over the past few decades has been more salient among married women than any other demographic groups (Greenwood, Guner and Vandenbroucke, 2017). Finally, this prominent change in participation also motivates us to focus on the extensive margin labor supply decision instead of the intensive margin time allocation decision in explaining the spatial divergence.

conditional on ζ^H and ζ^W, households solve the following problem:

$$\tilde{V}_d^{e,e'}(\zeta^H, \zeta^W)$$
$$= \delta^{e,e'} + \max_{H,W} \Big\{ \overbrace{\max_{h,n} \log\left(A_d^{e,e'}(I_{d,W}^{e,e'} - r_d h - p_n n)^{(1-\alpha-\beta)} h^\alpha n^\beta\right) + \xi_W}^{\text{Work}}, \overbrace{\max_{h} \log\left(A_d^{e,e'}(I_{d,H}^{e,e'} - r_d h)^{(1-\alpha)} h^\alpha (\bar{n}^{e'})^\beta\right) + \xi_H}^{\text{Housewife}} \Big\}$$
$$= \delta^{e,e'} + \log(A_d^{e,e'}) - \alpha \log(r_d) + \max\{c_1 + \log(I_{d,W}^{e,e'}) - \beta \log(p_n) + \xi_W,\ c_2 + \log(I_{d,H}^{e,e'}) + \beta \log(\bar{n}^{e'}) + \xi_H\},$$

where $c_1 = \log\left((1-\alpha-\beta)^{1-\alpha-\beta}\alpha^\alpha\beta^\beta\right)$, and $c_2 = \log\left((1-\alpha)^{1-\alpha}\alpha^\alpha\right)$.

The first component in household utility, $\delta^{e,e'}$, is a non-economic utility component, or 'love,' in a marriage of type (e,e'). Among other things, $\delta^{e,e'}$ could capture the affection between the couple or the pleasure from having one's own children. The second component in the first line of the equation reflects a choice of the household over whether the wife should **W**ork (W) or be a **H**ousewife (H). Working increases the income of the household ($I_{d,W}^{e,e'} > I_{d,H}^{e,e'}$), but staying at home adds home goods to the utility, reflected in an exogenous parameter $\bar{n}^{e'}$. Beyond the tangible services provided by the wife via home production, $\bar{n}^{e'}$ can also be more broadly interpreted as the utility from succumbing to the social prejudice against working wives. Finally, the value of amenities for the household is denoted by $A_d^{e,e'}$. It appears natural to assume that $A_d^{e,e'}$ depends on the amenities of both spouses, so we specify the following:

$$A_d^{e,e'} = (A_d^e)^{\frac{1}{2}}(A_d^{e'})^{\frac{1}{2}}.$$

Under this assumption, same-skill couples appreciate local amenities in the same way as if they were singles; mixed-skill couples, on the other hand, will have a more skill-balanced appreciation for different amenities. We will explore the impact of this assumption in Section 5.

From the properties of the Gumble distribution, the expected value of the household utility across the realizations of ζ^H, ζ^W draws is:

$$\overline{V}_d^{e,e'} = \mathbb{E}\tilde{V}_d^{e,e'}(\zeta^H, \zeta^W) \tag{14}$$
$$= \delta^{e,e'} + \log(A_d^{e,e'}) - \alpha \log(r_d) + \frac{\bar{\gamma}}{\eta_F^{e'}} + \frac{1}{\eta_F^{e'}} \log\left([\exp(c_1) \cdot I_{d,W}^{e,e'} \cdot p_h^{-\beta}]^{\eta_F^{e'}} + [\exp(c_2) \cdot I_{d,H}^{e,e'} \cdot (\bar{n}^{e'})^\beta]^{\eta_F^{e'}}\right),$$

in which the household income $I_{d,X}^{e,e'}$, $X \in \{H, W\}$ (for home and work) is

$$I_{d,X}^{e,e'} = \begin{cases} W_{d,M}^e + W_{d,F}^{e'} + 2t, & \text{if } X = W. \\ W_{d,M}^e + 2t, & \text{if } X = H. \end{cases}$$

The share of households of type (e,e') with a working wife is

$$l_d^{e,e'} = \frac{[\exp(c_1) \cdot I_{d,W}^{e,e'} \cdot p_h^{-\beta}]^{\eta_F^{e'}}}{[\exp(c_1) \cdot I_{d,W}^{e,e'} \cdot p_h^{-\beta}]^{\eta_F^{e'}} + [\exp(c_2) \cdot I_{d,H}^{e,e'} \cdot (\bar{n}^{e'})^\beta]^{\eta_F^{e'}}}. \tag{15}$$

3.4 Agglomeration, Congestion, and Housing

So far, we have described individual and household decisions, taking wages, amenities, and housing prices as given. We now develop the general equilibrium structure that determines these outcomes.

Labor supply and production. Given individuals' migration choices, the number of *residents* with gender s and skill e in location d, denoted by $\tilde{N}^e_{d,s}$, is:

$$\begin{aligned} \tilde{N}^e_{d,M} &= \sum_{o=1}^{S} [L^e_{o,M} \cdot \pi^e_{od,M}] \\ \tilde{N}^e_{d,F} &= \sum_{o=1}^{S} [L^e_{o,F} \cdot \pi^e_{od,F}], \end{aligned} \tag{16}$$

which is the sum over the people moving to d from different cities.

All men and single women work; among the married women, some stay at home, so the total *employment* in city d of a given demographic group is:

$$\begin{aligned} N^e_{d,M} &= \tilde{N}^e_{d,M} \\ N^e_{d,F} &= \sum_{o=1}^{S} [\tilde{N}^e_{d,F} \cdot (r^{H,e}_{d,F} \cdot l^{H,e}_d + r^{L,e}_{d,F} \cdot l^{L,e}_d + r^{\varnothing,e}_{d,F})]. \end{aligned} \tag{17}$$

Cities differ in their productivity, and the same number of workers can be turned into more units of effective labor in more productive cities. Let E^H_d and E^L_d be the effective labor supply of high- and low-skill workers in city d, receptively, we assume

$$\begin{aligned} E^H_d &= K^H_d (N^H_{d,M} + \beta^H_F N^H_{d,F}) \\ E^L_d &= K^L_d (N^L_{d,M} + \beta^L_F N^L_{d,F}). \end{aligned} \tag{18}$$

In this equation, $\beta^e_F \leq 1$, for $e \in \{H, L\}$ captures the productivity impact of the gender bias in workplace for women—because of the bias, female workers are not allocated efficiently and their productivity hindered. We will discipline β^e_F using the gender wage gap data. K^H_d and K^L_d are skill-specific productivity for the two types of workers in city d, which depends on both the exogenous characteristics of a city and the endogenous spillovers among workers. Following Fajgelbaum and Gaubert (2020), we specify:

$$\begin{aligned} K^H_d &= \bar{K}^H_d \cdot (N^H_{d,M} + N^H_{d,F})^{\gamma_{H,H}} \cdot (N^L_{d,M} + N^L_{d,F})^{\gamma_{L,H}} \\ K^L_d &= \bar{K}^L_d \cdot (N^H_{d,M} + N^H_{d,F})^{\gamma_{H,L}} \cdot (N^L_{d,M} + N^L_{d,F})^{\gamma_{L,L}}, \end{aligned} \tag{19}$$

in which $\bar{K}^H_d$ and $\bar{K}^L_d$ are skill-specific exogenous components of city d productivity. $\gamma_{e,e'}$ is the spillover effect from skill e to skill e'. By writing the spillovers as a function of employment rather than of population, we postulate that it is the interaction in the workplace, as opposed to the size of a city per se, that generates the spillovers.

High- and low-skill labor are imperfect substitutes and they are combined into the composite labor using the following constant elasticity of substitution technology.

$$E_d = \left((E_d^H)^\rho + (E_d^L)^\rho\right)^{\frac{1}{\rho}}. \tag{20}$$

Effective labor is used in the production of housing and the consumption good. The consumption good is the numeraire of the economy and is produced with the following technology:

$$Y_d = E_d^C, \tag{21}$$

where E_d^C is the amount of effective labor allocated to consumption good production. This numeraire consumption good can be converted to market alternatives of home goods n with a converstion rate of p_n. The remaining $E_d - E_d^C$ labor will be used in producing housing. The wage for each type of workers (denominated in the consumption good) is

$$\begin{aligned} W_{d,M}^H &= K_d^H \cdot E_d^{1-\rho}(E_d^H)^{\rho-1}; \qquad W_{d,F}^H = K_d^H \cdot \beta_F^H \cdot E_d^{1-\rho}(E_d^H)^{\rho-1} \\ W_{d,M}^L &= K_d^L \cdot E_d^{1-\rho}(E_d^l)^{\rho-1}; \qquad W_{d,F}^L = K_d^L \cdot \beta_F^L \cdot E_d^{1-\rho}(E_d^L)^{\rho-1}. \end{aligned} \tag{22}$$

Note that wages are determined by both the fundamental productivity and endogenous agglomeration forces. However, the spillovers among workers are not internalized and therefore do not directly enter wages.

Amenities. The amenities of city d are

$$\begin{aligned} A_d^H &= \bar{A}_d^H \cdot (\tilde{N}_{d,M}^H + \tilde{N}_{d,F}^H)^{\sigma_{H,H}} \cdot (\tilde{N}_{d,M}^L + \tilde{N}_{d,F}^L)^{\sigma_{L,H}} \\ A_d^L &= \bar{A}_d^L \cdot (\tilde{N}_{d,M}^H + \tilde{N}_{d,F}^H)^{\sigma_{H,L}} \cdot (\tilde{N}_{d,M}^L + \tilde{N}_{d,F}^L)^{\sigma_{L,L}}, \end{aligned} \tag{23}$$

in which $\bar{A}_d^e$ is the exogenous amenities component of city d and the remaining of the equation is the endogenous component. In particular, $\sigma_{e,e'}$ is the spillover in amenities from people with skill e to those with e'. Different from the spillovers in the labor market, which take place among workers, spillovers in amenities are functions of the number of residents. This captures that it is the size of the market for the amenities—restaurants, parks, museums—rather than employment, that drives these spillovers.

Housing. Housing is produced using labor under a decreasing return to scale function:

$$H_d = \bar{H}_d(E_d - E_d^C)^{\epsilon_d}.$$

$\bar{H}_d$ captures the overall level of land supply in city d and $0 < \epsilon_d < 1$ captures the difficulty in building additional housing due to the local zoning restrictions or topographical conditions. Representative housing producer takes rent as given and chooses the supply, giving us the local housing supply curve:

$$\log(H_d) = \frac{\epsilon_d}{1-\epsilon_d} \cdot \log(\epsilon_d \bar{H}_d^{\frac{1}{\epsilon_d}}) + \frac{\epsilon_d}{1-\epsilon_d} \cdot \log(r_d) \tag{24}$$

in which $\frac{\epsilon_d}{1-\epsilon_d}$ is the supply elasticity and $\frac{\epsilon_d}{1-\epsilon_d} \cdot \log(\epsilon_d \bar{H}_d^{\frac{1}{\epsilon_d}})$ the supply shifter.

The housing market clearing conditions is

$$H_d = \frac{\alpha(E_d + t \cdot \sum_{s,e} \tilde{N}^e_{d,s})}{r_d}, \tag{25}$$

where α is the housing share of consumption, E_d is the total labor income to residents of city d, and t is the lump-sum transfer to residents.

These transfers are financed by the profit from the housing sector (or equivalently, the payment to land as the fixed input in housing production). Budget balance of transfers implies that total transfer equals the total profits from land ($1-\epsilon_d$ share of housing expenditures). We have:

$$t = \frac{\sum_d (1-\epsilon_d) \cdot H_d \cdot r_d}{\sum_{o,s,e} L^e_{o,s}}. \tag{26}$$

3.5 Definition of the Equilibrium

The fundamental parameters of the model are: the exogenous components of city productivity and amenities $\{\bar{A}^e_d,\ \bar{K}^e_d\}$, local land supply shifter and elasticity, $\{\bar{H}_d,\ \epsilon_d\}$, the endowments of the four types of workers in each city $\{L^e_{o,s}\}$, the non-economic component of marriage surplus $\{\delta^{e,e'}\}$, the parameters governing idiosyncratic taste draws $\{\kappa^e_s, \theta^e_s, \eta^e_F\}$, gender biases and home production value $\{\beta^e_F, \bar{n}^e\}$, migration costs $\{d^e_{od,s}\}$, spillover elasticities $\{\gamma_{e,e'},\ \sigma_{e,e'}\}$, and Cobb-Douglas shares in the utility function $\alpha,\ \beta$.

Definition 2. *A competitive equilibrium of the model is defined as a set of prices and allocations such that:*

1. *Given the equilibrium wages and prices, individual and household choices are optimal, and the marriage market in each city is a stable match:*
   - *(a) The expected utility for households, $\overline{V}^{e,\varnothing}_d$, $\overline{V}^{\varnothing,e'}_d$, and $\overline{V}^{e,e'}_d$ are given by equations (13) and (14); the female labor force participation decision, $l^{e,e'}_d$, is given by equation (15).*
   - *(b) Given the expected household utility, each local marriage markets is a stable match, characterized by equations (5), (8), and (9). The implied expected utility before the marriage market is given by equation (10).*
   - *(c) Workers make optimal migration decisions according to equation (2).*
2. *Under the optimal individual and household decisions, the resulting allocation of workers are consistent with general equilibrium outcomes:*
   - *(a) The number of people working and residing in location d are given by equations (16) and (17).*

(b) *Wages are given by equations (19) and (22).*

(c) *Amenities are given by equation (23).*

(d) *Housing markets clear, i.e., equations (24) and (25).*

(e) *The transfer from the rent of land is given by equation (26)*

3.6 Marriage and Implications for the Spatial Economy

In our model, marriages embody three main ingredients. First, couples have access to a technology that singles do not have—home production. This technology provides an option value to households through idiosyncratic draws even when it is not being used. Second, couples enjoy local amenities differently from the singles. Finally, marriages have a love component, $\delta^{e,e'}$, which also shapes the sorting in the marriage market.

Changes in these three ingredients determine the marriage surplus and location choices. For example, a reduction in the discrimination against women in the workplace and at home (through an increase in β_F^e or a decrease in $\bar{n}^e$) encourages women to participate in the labor market. This change in turn decreases the option value from home production for married couples, thus weakening the importance of marriage in location choices but strengthening the importance of the labor market returns. Similarly, given that cities differ in their marriage market outcomes, a shift in the social norm on marriages captured in $\delta^{e,e'}$ can also alter individuals' choice of cities. We will discipline the model with data to uncover the strength of these forces.

4 Parameterization

In this section, we parameterize the model to match the U.S. economy in year 2000. Some of our model elasticities appear in workhorse quantitative spatial models and have been estimated in the literature; others can be estimated directly in structural equations independent of the rest of model. In Sections 4.1 to 4.4, we explain how we estimate or choose the values for these parameters. We then describe how the remaining model parameters are pinned down jointly in Section 4.5 and demonstrate how the model fits the data in Section 4.6.

4.1 Parameters Assigned Directly

Spillover and substitution between skills. $\{\gamma_{e,e'}, \sigma_{e,e'}\}$ are the spillovers in productivity and amenities, and have been estimated in a voluminous urban economics literature. Most estimates do not consider the asymmetric spillovers between high- and low-skill people. In recent work, Fajgelbaum and Gaubert (2020) convert existing estimates to be consistent with the asymmetric setup that we adopt in this paper. We take their values as below:

$$\sigma_{H,H} = 0.77,\ \sigma_{H,L} = 0.18,\ \sigma_{L,H} = -1.24,\ \sigma_{L,L} = -0.43;$$
$$\gamma_{H,H} = 0.05,\ \gamma_{H,L} = 0.04,\ \gamma_{L,H} = 0.02,\ \gamma_{L,L} = 0.003.$$

These values suggest a higher importance of spillovers via amenities than via productivity; it also highlights an asymmetry: while high-skill workers might generate positive amenity spillovers to low-skill workers, the reverse spillovers are negative.

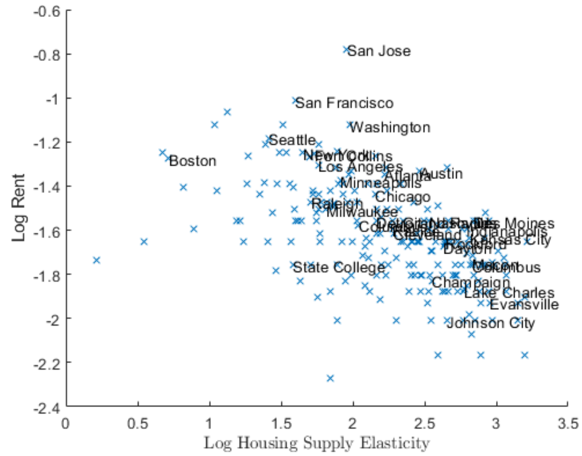


Figure 2: Rent and Housing Supply Elasticity

Note: This figure plots the log housing supply elasticities and the log of average rents of cities. Rent is from the 2000 census; housing supply elasticities are calculated from the estimates of Diamond (2016).

Parameter ρ governs the elasticity of substitution between high- and low-skill workers. Following Fajgelbaum and Gaubert (2020), we set $\rho = 0.392$.

Housing supply elasticities. Our model allows cities to differ in housing supply elasticity $\frac{\epsilon_d}{1-\epsilon_d}$. Diamond (2016) estimates ϵ_d as a function of several local topological features assembled in Saiz (2010). We use these estimates to generate city-specific housing supply elasticities. Figure 2 plots the logarithm of these elasticities against the logarithm of city rents calculated from the 2000 Census. It shows that where housing supply is more elastic, rent tends to be lower.

Migration elasticities. The parameters governing the migration elasticities θ_s^e differ by skill. Using the U.S. data and instrumenting wage income with shift-share instruments, Diamond (2016) recovers the income elasticities of migration separately for high- and low-skill workers, which correspond to θ_s^e in our model. We adopt the following estimates of hers: $\theta_F^H = \theta_M^H = 4.98$, $\theta_F^L = \theta_M^L = 3.26$.

Housing and home-production shares. To determine the housing share of consumption α, we use the 5% sample of the 2000 census (Ruggles et al., 2021) to calculate the rent share of income for renters.[16] The sample average of the rent share gives us $\alpha = 0.25$. Parameter β governs the importance of home-produced goods in utility. We treat staying-at home wives as having a full time job. Assuming 40 hours of work per week, this job takes about one-fifth of all non-sleep hours between the spouses. We thus set $\beta = 0.2$.[17] With Cobb-Douglas preference, the price of home goods does not affect the consumption allocation, so we normalize $p_n = 1$ in the baseline economy.

The gender wage gap. We calculate the gender wage gap as the gender dummy in an

[16] We use the rent share among renters instead of the housing expenditure share among homeowners (from e.g., the Consumer Expenditure Survey) because part of the latter is equity investment, instead of consumption.

[17] We will use $\bar{n}^e$ to match the female labor force participation. Different values of β (and p_n) will lead to differences in inferred $\bar{n}^e$, but conditional on that, the results from the counterfactual of eliminating marriages will not be affected. See Appendix Table C.1 for results showing the insensitivity of the model predictions to the choice of β.

individual-level Mincer regression, controlling for age, education, and location using the 2000 data. This procedure results in $\beta_F^H = 0.76$, $\beta_F^L = 0.74$.

4.2 Female Labor Force Participation

The remaining parameters do not have theory-consistent estimates readily available, so we estimate them internally. To estimate the parameter governing the supply elasticity of female labor, we arrange equation (15) to obtain:

$$\log(\frac{l_d^{e,e'}}{1 - l_d^{e,e'}}) = \eta_F^{e'} \cdot [\text{constant} - \beta \log(p_n) - \beta \log(\bar{n}^{e'})] + \eta_F^{e'} \cdot [\log(I_{d,W}^{e,e'}) - \log(I_{d,H}^{e,e'})].$$

In the outcome variable, $l_d^{e,e'}$, is the fraction of wives in (e, e') marriages in city d who work. On the right hand side of the equation, $I_{d,W}^{e,e'}$ is the household income if the wife works, while $I_{d,H}^{e,e'}$ is the household income if the wife stays at home; $\bar{n}^{e'}$ is the home good provided by staying-at-home wives with skill e'; p_n is a measure of the market price of household products. The parameter of interest is $\eta_F^{e'}$, which is the response in the log odds ratio to the percentage increase in household income from the wife's participation in the labor market.

In the right hand side of the specification, $I_{d,W}^{e,e'}$ and $I_{d,H}^{e,e'}$ can be directly measured as the average annual income of different household types, while p_n and $\bar{n}^{e'}$ are not observed. In principle, given the assumption that these two parameters are common to all cities, they will be absorbed in the constant term. However, it is possible that cities differ in the market price of home goods or the household demand for these goods (due to the differences in the composition of household types). To address this concern, we use two variables to capture the regional variation in p_n: log median rental price and the share of low-skill immigrants in work-age population; we also include proxies for the demand for home goods: the age distribution of type (e, e') households in city d and the share among them with young children. Those proxy variables are denoted as $\hat{p}_n$ and $\hat{n}^{e'}$.

We separately estimate η_F^H and η_F^L for households with high- and low-skill wives. Formally, our regression specification is :

$$\log(\frac{l_d^{e,e'}}{1 - l_d^{e,e'}}) = \eta_F^{e'} \cdot [\log(I_{d,W}^{e,e'}) - \log(I_{d,H}^{e,e'})] + f(\hat{p}_n, \hat{n}^{e'}) + \lambda^{e,e'} + \varepsilon_d^{e,e'}. \tag{27}$$

$\lambda^{e,e'}$ is the fixed effect for (e, e') households, $f(\hat{p}_n, \hat{n}^{e'})$ is a flexible function of proxy variables, and $\varepsilon_d^{e,e'}$ is the measurement error. Each observation is a marriage type by city. Our estimation uses the 5% sample of the 2000 Census, restricted to married couples with both spouses present and at least one of them between 40 and 54 years old. This age restriction ensures that we focus on the 'middle-aged' labor supply decisions among people who are generally past the age for the (initial) marriage market but have not yet reached the age for retirement.

Table 1 reports the estimation results. The labor supply response $\eta_F^{e'}$ is 0.87 for high-skill women and 3.33 for low-skill women. With the dependent variable being $\log(l_d^{e,e'}/(1 - l_d^{e,e'}))$, $\eta_F^{e'}$

Table 1: Estimation of Female Labor Force Participation Parameter: $\eta_F^{e'}$

	(1)	(2)
	dep var: $\log(\frac{l_d^{e,e'}}{1-l_d^{e,e'}})$	
	$e' = H$	$e' = L$
$\log(I_{d,W}^{e,e'}) - \log(I_{d,H}^{e,e'})$	0.866	3.334
	(0.407)	(0.400)
Controls		
log rent	X	X
% having young children	X	X
distr. of husband age	X	X
Household type FE	X	X
	$(H,H),(L,H)$	$(H,L),(L,L)$
N	653	654

Note: Estimation of equation (27) using the 2000 Census. Each observation is a city-marriage type combination. The dependent variable is the log odds ratio for female labor force participation, for high-skill women in Column 1 and low-skill women in Column 2. Robust standard errors are in parentheses.

is not the female labor force participation elasticity as usually defined. Converting our estimate to the conventionally defined female labor force participation elasticity—where $\log(l_d^{e,e'})$ instead of the log odds ratio is used as the dependent variable—gives us an elasticity of 0.26 for high-skill women and 0.97 for low-skill women. These values are in line with existing estimates. For example, using the Current Population Survey between 1999 and 2001, Blau and Kahn (2007) estimate an elasticity of 0.26 for high-skill women and an elasticity of 0.6 for low-skill women.

4.3 Marriage Preference Heterogeneity

Parameters $\{\kappa_s^e\}$ determine the importance for an individual of type (s,e) of their personal tastes for partners. We estimate these parameters by exploiting the difference across cities in local prices and the composition of marriages.

Let $q_d^{e,e'}$ be the number of type (e,e') marriages in city d, $q_d^{e,\varnothing}$ be the number of single men with skill e, and $q_d^{\varnothing,e'}$ be the number of single women with skill e' in city d. Equations (8) and (9) imply the following (see the Appendix for the derivation):

$$q_d^{e,e'} = [q_d^{e,\varnothing} \cdot \frac{\exp(\kappa_M^e U_{d,M}^{e,e'})}{\exp(\kappa_M^e U_{d,M}^{e,\varnothing})}]^{\frac{\frac{1}{\kappa_M^e}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}}} \cdot [q_d^{\varnothing,e'} \cdot \frac{\exp(\kappa_F^{e'} U_{d,F}^{e,e'})}{\exp(\kappa_F^{e'} U_{d,F}^{\varnothing,e'})}]^{\frac{\frac{1}{\kappa_F^{e'}}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}}}.$$

Taking logarithm on both sides and using Lemma 1 gives us:

$$\log(q_d^{e,e'}) = \frac{1}{\frac{1}{\kappa_M^e} + \frac{1}{\kappa_F^{e'}}} \cdot \underbrace{(\overline{V}_d^{e,e'} - \overline{V}_d^{e,\varnothing} - \overline{V}_d^{\varnothing,e'})}_{\text{systematic marriage surplus}} + \frac{\frac{1}{\kappa_M^e}}{\frac{1}{\kappa_M^e} + \frac{1}{\kappa_F^{e'}}} \log(q_d^{e,\varnothing}) + \frac{\frac{1}{\kappa_F^{e'}}}{\frac{1}{\kappa_M^e} + \frac{1}{\kappa_F^{e'}}} \log(q_d^{\varnothing,e'}). \tag{28}$$

This equation expresses the number of type (e, e') matches in city d as a function of the marriage surplus in city d and the number of single persons in these two types. Intuitively, if the number of marriages respond more strongly to the marriage surplus, it means that the idisocynratic draws are relatively less important for marriage market outcomes, i.e., κ_F^e and $\kappa_F^{e'}$ must be large. Therefore, if we can measure the marriage surplus, the variation in such surplus across cities identifies $1/(\frac{1}{\kappa_M^e} + \frac{1}{\kappa_F^{e'}})$. Conditional on $1/(\frac{1}{\kappa_M^e} + \frac{1}{\kappa_F^{e'}})$, the numbers of single (M, e) and (F, e') types in the city also matter for $q^{e,e'}$, which enables separate identification of κ_M^e and $\kappa_F^{e'}$.

Although the marriage surplus is not directly observable, from equations (13), (14), and (15), it can be expressed as a function of measurable outcomes, as below:

$$\begin{aligned}\overline{V}_d^{e,e'} - \overline{V}_d^{e,\varnothing} - \overline{V}_d^{\varnothing,e'} = &\overbrace{[\delta^{e,e'} + \frac{\bar{\gamma}}{\theta_F^{e'}} - c_1 + \beta \log(p_n)]}^{(e,e')\text{ specific constant}} + [\log(A_d^{e,e'}) - \log(A_d^e) - \log(A_d^{e'})] \\ &+ [\log(\frac{I_d^{e,e'}}{I_{d,M}^e \cdot I_{d,F}^{e'}}) - \frac{1}{\eta_F^{e'}} \log(l_d^{e,e'})] + \alpha \log(r_d).\end{aligned}$$

This equation decomposes the marriage surplus into the sum of four terms: a constant that is specific to each type of marriage but common across cities; the difference between the household appreciation of amenities and the sum of the two spouses' own appreciation of the amenities, which can be non-parametrically controlled for through various proxies for amenities; a term on the income of married couples relative to that of the two spouses as singles, adjusted for the labor force participation rate; finally, the log rent of city d enters this with a coefficient α, the share of expenditure on rent.

Plugging the above into equation (28) and collecting the equation for all four types of marriages, we have the following set of equations:

$$\begin{aligned}\log(q_d^{H,H}) &= g(\mathbf{X}_d^{H,H}) + \frac{1}{\frac{1}{\kappa_M^H} + \frac{1}{\kappa_F^H}}[\alpha \log(r_d) + \frac{1}{\kappa_M^H}\log(q_d^{H,\varnothing}) + \frac{1}{\kappa_F^H}\log(q_d^{\varnothing,H})] + \varepsilon_d^{H,H} \\ \log(q_d^{H,L}) &= g(\mathbf{X}_d^{H,L}) + \frac{1}{\frac{1}{\kappa_M^H} + \frac{1}{\kappa_F^L}}[\alpha \log(r_d) + \frac{1}{\kappa_M^H}\log(q_d^{H,\varnothing}) + \frac{1}{\kappa_F^L}\log(q_d^{\varnothing,L})] + \varepsilon_d^{H,L} \\ \log(q_d^{L,H}) &= g(\mathbf{X}_d^{L,H}) + \frac{1}{\frac{1}{\kappa_M^L} + \frac{1}{\kappa_F^H}}[\alpha \log(r_d) + \frac{1}{\kappa_M^L}\log(q_d^{L,\varnothing}) + \frac{1}{\kappa_F^H}\log(q_d^{\varnothing,H})] + \varepsilon_d^{L,H} \\ \log(q_d^{L,L}) &= g(\mathbf{X}_d^{L,L}) + \frac{1}{\frac{1}{\kappa_M^L} + \frac{1}{\kappa_F^L}}[\alpha \log(r_d) + \frac{1}{\kappa_M^L}\log(q_d^{L,\varnothing}) + \frac{1}{\kappa_F^L}\log(q_d^{\varnothing,L})] + \varepsilon_d^{L,L}.\end{aligned} \tag{29}$$

In equation (29), we capture a number of parametric and non-parametric controls in $g(\mathbf{X}_d^{e,e'})$. $\mathbf{X}_d^{e,e'} = \{\lambda^{e,e'},\ \hat{A}_d,\ \hat{p}_n,\ I_d^{e,e'},\ \log(\frac{I_d^{e,e'}}{I_{d,M}^{e} \cdot I_{d,F}^{e'}})\}$, in which $\lambda^{e,e'}$ is a fixed effect for marriage type (e, e') and $\hat{A}_d$ denotes a flexible function of common proxies for amenities, including climate (annual precipitation, average daily low temperature in January) and the presence of services (number of restaurants, entertainment businesses, health care facilities per capita). We also account for possible differences in the cost of home goods across cities by controlling for a flexible function of log median rental price and the share of low-skill immigrants in work-age population, denoted by $\hat{p}_n$. Last but not least, we introduce a measurement error $\varepsilon_d^{e,e'}$ in each equation.

This set of equations resembles the equilibrium relationship of transferable utility models derived in Graham (2013). In existing empirical studies of such models (Choo and Siow, 2006; Chiappori et al., 2017), researchers often assume the dispersion parameters governing individuals' tastes to be the same across groups and then normalize one parameter to 1. Indeed, when the only available data are the marriage patterns in *one* market, these dispersion parameters are not separately identifiable from the marriage surplus terms. In our case, the spatial dimension of the model is key for identifying all four dispersion parameters. First, through individuals' location choices, the model links an observable component of the location-specific utility, $\log(r_d)$, to marriage patterns. Taking $\alpha = 0.25$, the variation in log rent across locations identifies $\frac{1}{\frac{1}{\kappa_M^e} + \frac{1}{\kappa_F^{e'}}}$. Second, by treating each city as a separate local marriage market, we exploit how the composition of marriages varies with the number of available single people of each type across cities to identify $\frac{1}{\kappa_M^e}$ and $\frac{1}{\kappa_F^{e'}}$.

We estimate equation (29) for κ_M^H, κ_M^L, κ_F^H, and κ_F^L. Because each of these parameters appears in two equations, the model imposes cross-equation restrictions on their values.[18] We estimate the four equations jointly. Our estimation focuses on the sample of people from the 2000 census aged between 40 and 54, the stage after the marriage market is largely settled.

Table 2 reports the results. Column 1 reports the baseline specification without controlling for amenities; column 2 controls for a flexible function of variables capturing the local climate; column 3, our preferred specification, further controls for the availability of various services. According to the result from column 3, high-skill men and women have similar degrees of heterogeneity in their preference for partner types; on the other hand, the idisocynratic preferences of low-skill women are more homogeneous than those of low-skill men.

4.4 Migration Cost Parameters

In addition to the income elasticities of migration θ_s^e, migration costs are also important for migration decisions. We specify the cost of moving from o to d for type (s, e) as:

[18]That is, there are twelve coefficients across the four equations, but these coefficients are uniquely determined by four structural parameters.

Table 2: Marriage Elasticities (κ_s^e)

	(1)	(2)	(3)
	dep var: $\log(q_d^{e,e'})$		
κ_M^H	2.37	1.84	1.62
	(0.35)	(0.37)	(0.40)
κ_F^H	2.13	1.64	1.66
	(0.40)	(0.39)	(0.47)
κ_M^L	1.21	0.87	0.71
	(0.27)	(0.24)	(0.24)
κ_F^L	5.13	3.53	2.61
	(1.39)	(1.08)	(0.91)
Controls			
$\log(\frac{I_d^{e,e'}}{I_{d,M}^e \cdot I_{d,F}^{e'}})$	X	X	X
$\log(l_d^{e,e'})$	X	X	X
$\hat{p}_n$	X	X	X
$\log(\hat{A}_d)$			
climate		X	X
services			X
N	1181	1181	1181

Note: Joint estimation of equations in (29). Data from 2000 Census. Each observation is a city-marriage type combination. The dependent variable is the log number of marriages by marriage type. Robust standard error in parentheses.

$$d_{od,s}^e = \mathbb{I}_{S(o) \neq S(d)} \cdot \sum_{b=1}^{5} \tau_{b,s}^e \cdot \mathbb{I}_b. \tag{30}$$

In the specification, $S(o)$ and $S(d)$ denote the state of the origin and destination city, respectively. Because the census data only include the birth state but not birth city, we assume migration costs are zero within a state. Between states, migration costs depend on the population weighted average distance between state $S(o)$ and MSA d. We divide distances into five bins, indicated by function $\mathbb{I}_b$: (1) within 160 km, (2) between 160 and 500 km, (3) between 500 and 1000 km, (4) between 1000 and 2000 km, and (5) more than 2000 km.

Combining this cost function with equation (2), we obtain the following estimation equation:

$$\log(N_{od,s}^e) = \lambda_{o,s}^e + \lambda_{d,s}^e - \theta_s^e \cdot \mathbb{I}_{S(o) \neq S(d)} \cdot \sum_{b=1}^{5} \tau_{b,s}^e \cdot \mathbb{I}_b + \varepsilon_{od,s}^e, \tag{31}$$

where $N_{od,s}^e$ is the number of individuals of gender $s \in \{M, F\}$ and skill level $e \in \{H, L\}$, who were born in state o and now live in MSA d. $\lambda_{o,s}^e$ and $\lambda_{d,s}^e$ are birth state and destination MSA fixed effects, capturing state o's endowment of labor and access to other cities ('multilateral

Table 3: Inter-State Migration Rate (%)

	1960		2000	
	skilled	unskilled	skilled	unskilled
men	45.1	31.7	47.7	30.4
women	46.2	32.7	45.2	30.9

Note: Data are from the 1960 and 2000 censuses. The sample includes native-born individuals between 25 and 39, currently not enrolled in school. Migration is defined as living outside the state of birth.

Table 4: Distance Elasticities of Migration Cost

	(1)	(2)	(3)	(4)
	male		female	
dep var: $\log(N^e_{od,s})$	high	low	high	low
$\theta^e_s \cdot \tau^e_{1,s}$	1.870	2.224	1.957	2.148
	(0.112)	(0.135)	(0.114)	(0.141)
$\theta^e_s \cdot \tau^e_{2,s}$	2.621	3.312	2.713	3.260
	(0.051)	(0.062)	(0.053)	(0.061)
$\theta^e_s \cdot \tau^e_{3,s}$	3.510	4.315	3.638	4.280
	(0.048)	(0.058)	(0.050)	(0.058)
$\theta^e_s \cdot \tau^e_{4,s}$	4.025	4.888	4.161	4.868
	(0.048)	(0.059)	(0.050)	(0.058)
$\theta^e_s \cdot \tau^e_{5,s}$	4.346	5.378	4.529	5.365
	(0.050)	(0.061)	(0.052)	(0.060)
destination MSA FE ($\lambda^e_{d,s}$)	X	X	X	X
state-of-origin FE ($\lambda^e_{o,s}$)	X	X	X	X
N	11099	13529	11436	13586

Note: Data are from the 2000 Census. The sample includes native-born individuals who reside in an MSA between ages 25 and 39. Each observation is a state-of-origin by destination MSA pair. The outcome variable is $\log(N^e_{od,s})$. Robust standard errors are in parentheses.

resistance'), and city d's expected utility, respectively.

To be consistent with the model, in which migration decisions are made by young adults who have not completed the matching process, our estimation of migration costs focuses on the sample of native-born individuals between 25 and 39 years old.[19] Table 3 shows the lifetime inter-state migration rate for both 2000 and 1960. In 2000, more than 45% of people with a college degree and 30% of people without a college degree lived outside of their birth states. For both groups, the migration rate is largely stable over the four decades.

Table 4 reports the estimation results using the 2000 census data. Migration is hindered by

[19]Note that in estimating the marital market outcomes in Section 4.3, we restrict the sample to those between 40 and 54 years old, our empirical analysis effectively splits adulthood into two stages: before ($\leq$ 39) and after ($\geq$ 40) the conclusion of the initial marriage market. The empirical patterns for migration are not materially different if we focus on younger cohorts, say, individuals between 22 and 29 years old, who have completed their education.

geographic distances for all four types of individuals. Within the same skill level, the estimated coefficients are similar between men and women. Within the same gender, the coefficients tend to be larger for low-skill workers. Because the structural interpretation of the coefficients for distance bins in equation (31) is $\theta_s^e \cdot \tau_{b,s}^e$, we can divide these estimates by θ_s^e to obtain the estimates for $\tau_{b,s}^e$. We can then plug these into (30) to calculate the bilateral migration costs. Recalling that $\theta_s^H = 4.98 > \theta_s^L = 3.26$, our estimation implies substantially higher migration costs for low-skill workers.

4.5 Calibrating the Remaining Parameters Jointly

Panels A and B of Table 5 summarize the values of parameters discussed thus far. The remaining parameters affect model outcomes through the general equilibrium. We choose these parameters jointly.

Specifically, parameters $\bar{H}_d$, $\bar{A}_d^e$, and $\bar{K}_d^e$ vary by city. We recover these parameters by matching the rent, the number of high- and low-skill workers, and their averages wages in each city. $\delta^{e,e'}$ captures the utility of married couples (e, e') that is not a part of the economic return. We identify $\delta^{e,e'}$ using the share of people in a marriage (68%) and the shares of the four types of marriages. Finally, we identify $\bar{n}^e$ from the labor force participation rate of high- and low-skill women.

Appendix C describes the implementation of the calibration. Panel C of Table 5 summarizes the parameters. We find positive and generally large values for $\delta^{H,H}$ and $\delta^{L,L}$, which captures that empirically, marriages tend to be assortative. Among the mixed-skill marriages, $\delta^{L,H}$ is negative but $\delta^{H,L}$ is positive. This stark difference could be due to the social stereotype of the husband as the breadwinner of the family, which reduces the number of marriages with a low-skill husband and a high-skill wife. We also find that the preference of high-skill women for staying at home is lower than that of low-skill women.

4.6 Model Validation

In this subsection, we compare the model's predictions on non-targeted outcomes with the data to show that the model captures the essence of the two markets we are studying.

The composition of skills by gender. By design, our calibration ensures that the skill share of each city matches the data. However, the *gender composition* of skills and how this composition varies across cities are equilibrium outcomes that depend on the interaction between the local marriage and labor markets. We first verify that the model matches these dimensions well. Figure 3a plots the relative skill of women versus men (log (H/L) among women minus that among men) against the log skill ratio of cities. The values range between -0.2 and 0.2 but are centered around zero, mirroring that in the aggregate data, the college shares of men and women are similar—marriage markets play the role of an 'equalizer' for the gender composition of skills across cities. More importantly, the model also shows that in skill intensive cities, women are comparatively less skilled than men. This could be due to the higher marriage market return in skill intensive cities for low-skill women, a hypothesis made by Edlund (2005) based on an empirical observation from Sweden. Figure 3b shows that in the data, both patterns are also true qualitatively.

Table 5: Summary of Parameters

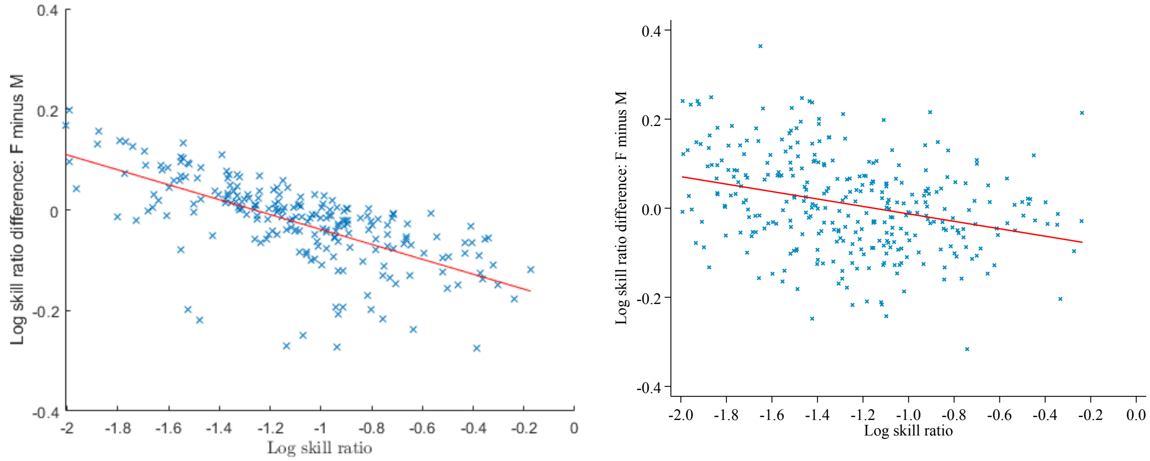
Parameters	Descriptions	Value	Targets/Source
A. Assigned directly			
$\sigma_{e,e'}$	amenity spillovers	$\sigma_{H,H} = 0.77$, $\sigma_{H,L} = 0.18$, $\sigma_{L,H} = -1.24$, $\sigma_{L,L} = -0.43$	Fajgelbaum and Gaubert (2020)
$\gamma_{e,e'}$	prod. spillovers	$\gamma_{H,H} = 0.05$, $\gamma_{H,L} = 0.04$, $\gamma_{L,H} = 0.02$, $\gamma_{L,L} = 0.003$	Fajgelbaum and Gaubert (2020)
ρ	substitution between skills	0.392	Fajgelbaum and Gaubert (2020)
β^e_F	gender wage gap	$\beta^H_F = 0.76, \beta^L_F = 0.74$	2000 Census
α	housing share	0.25	2000 Census
β	home-good share	0.2	2000 Census
ϵ_d	housing supply elast.	Figure 2	Diamond (2016)
θ^e_s	income elast. of migration	$\theta^H_M = \theta^H_F = 4.98$, $\theta^L_M = \theta^L_F = 3.26$	Diamond (2016)
B. Estimated independently			
η^e_F	labor force participation.	Table 1	2000 Census
κ^e_s	marriage taste shock	Table 2, column 3	2000 Census
$\tau^e_{b,s}$	migration cost	Table 4	2000 Census
C. Calibrated jointly			
$\bar{H}_d$	housing supply shifter	-	rent by city
$\bar{A}^e_d$	fund. amenities	-	emp by city × skill
$\bar{K}^e_d$	fund. prod.	-	wage by city × skill
$\delta^{e,e'}$	love	$\delta^{H,H} = 1.07$, $\delta^{H,L} = 0.28$, $\delta^{L,H} = -1.96$, $\delta^{L,L} = 1.20$	68% people in marriages; composition: 21% (H,H), 13% (H,L), 9% (L,H), 56% (L,L)
$\bar{n}^e$	home prod. pref.	$\bar{n}^H = 0.004$, $\bar{n}^L = 1.03$	labor force participation (83% and 73%)

Note: This table summarizes the values of parameters taken externally (Panel A), estimated by the authors independent of the rest of the model (Panel B), and calibrated jointly in equilibrium (Panel C).

The composition of households. Our calibration matches the shares of the four types of marriages in the aggregate, but does not impose how these shares should vary by city. Figure 4a plots the model's prediction on the composition of marriages in each city. As the skill intensity of a city increases, the share of (L, L) marriages in all local marriages decreases and the share of (H, H) marriages increases. The shares of mixed-skilled marriages also increase, with the share of the (H, L) type increasing at a faster rate than that of the (L, H) type. These patterns are broadly consistent with the data, shown in Figure 4b.

Figure 4c plots the composition of single households in each city. As the skill intensity of a city increases, not surprisingly, high-skill singles account for a larger fraction of local single households. Interestingly, even though in the aggregate, the fraction of men without a college degree are similar to the fraction of women without a college degree, across the city skill intensity spectrum, low-skill men account for a higher fraction of single households than low-skill women. Figure 4d plots the empirical counterparts.[20] It shows that the model is consistent with the data

[20]Since the static model does not incorporate divorce decisions, in counting single people in the data for Figures



(a) Log Relative Skill Ratio: Model

(b) Log Relative Skill Ratio: Data

Figure 3: Gender and Skill Compositions: Model versus Data

Note: The left panel plots $\log(\frac{\text{\# of high skill women}}{\text{\# of low skill women}})$ - $\log(\frac{\text{\# of high skill men}}{\text{\# of low skill men}})$ against $\log(\frac{\text{\# of high-skill people}}{\text{\# of low-skill people}})$; the right panels shows the empirical counterpart of this relationship.

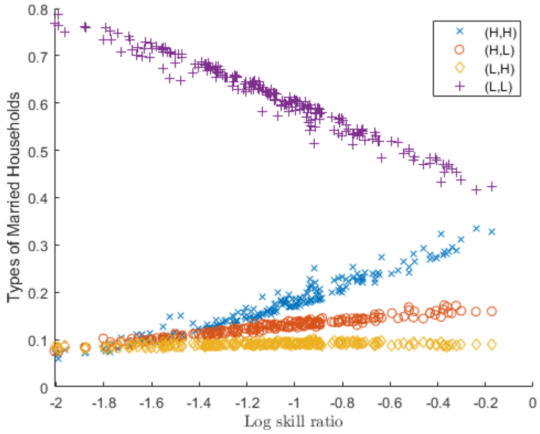
in both the overall share of each type among the singles and the relationship between these shares and the city skill intensity.

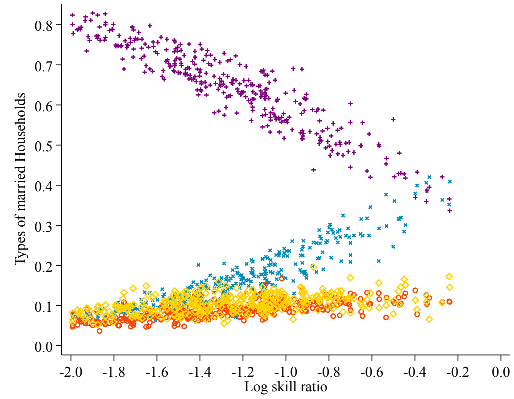
Lastly, we examine the single rate of cities. Figures 4e and 4f plot single rates by group. In the data, the average single rate across groups and cities is around 25%. The model predicts on overall a similar single rate and replicates the higher single rates among low-skill men and women seen in the data. Importantly, both figures show a noisy but positive correlation between single rates and city skill intensity for all four types of people.[21]

In summary, while not calibrated to these patterns, the model is able to produce salient features on variation in gender skill composition and marriage market outcomes across cities with different skill intensity. Importantly, in doing so, we have specified the key ingredients of the model as common across cities and let the equilibrium forces discipline how many people choose to marry and how. An alternative approach would be to allow for some of the model components to differ across cities. For example, one can specify the 'love' component as a function of skill ratio of a city, i.e., $\delta_d^{e,e'} = \delta^{e,e'} \cdot (\text{skill ratio})_d^{\Delta}$, which can be interpreted as some cities offering amenities that appeal more to married couples than singles—e.g. more accessible and less expensive daycare centers relative to restaurants and bars. One can then (1) either choose Δ to match the residual variation in marriage patterns across cities that our model fails to explain, or (2) to estimate Δ externally using proxies for such amenities. Since our model captures the empirical

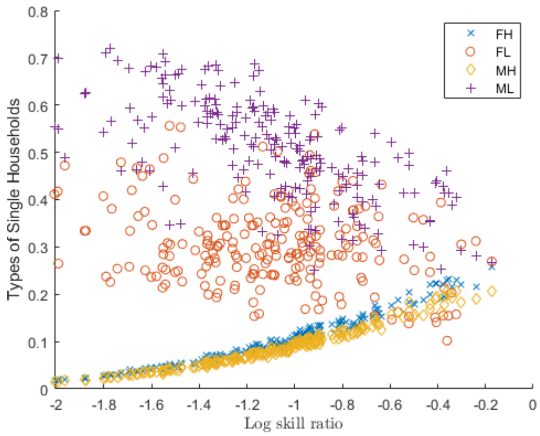
4d and 4f, we include people who are currently not in a marriage, effectively counting only people in 'successful' marriages as being married. Focusing only the never married generates similar patterns.

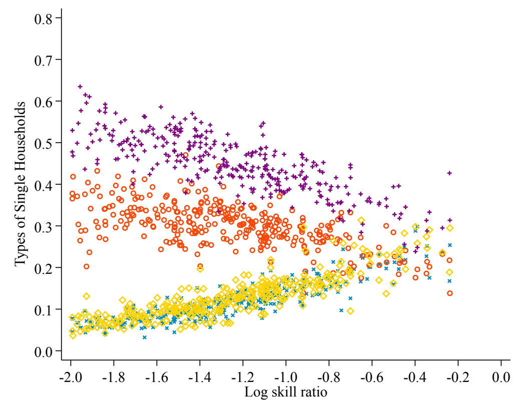
[21]One caveat in interpreting Figure 4f is that in the data, large and skill intensive cities might have higher shares of young people. To the extent that young people are less likely to be married, the pattern in Figure 4f could be driven by the age differences between cities. To address this concern, Appendix Figure A.2 plots the age-adjusted single rate, in which we regress whether a person is single on the skill intensity of his/her city, controlling for age-by-race fixed effects. The general upward trend there is as evident as—if not more evident than—in Figure 4f.



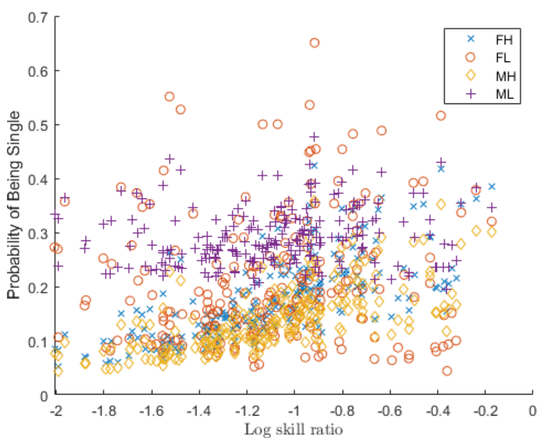


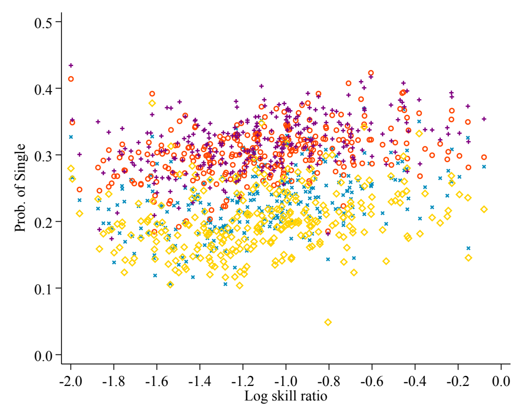
(a) Composition of Married Households: Model (b) Composition of Married Households: Data





(c) Composition of Single Households: Model (d) Composition of Single Households: Data





(e) Probability of Being Single: Model (f) Probability of Being Single: Data

Figure 4: The Composition of Households: Model versus Data

Note: the left panel shows the model's predictions on the composition of four types of marriages (panel a), the composition of single people (panel c), and the single rate (panel e) of cities; the right panels show the empirical counterparts of the left panels.

variation in marriage outcomes well, the first approach will likely find a minor role for Δ. In Appendix Figure A.3, we construct proxies for differential amenities for married and singles and find that the second approach is also unlikely to undermine the importance of our mechanisms.

5 The Impacts of Marriage on the Spatial Distribution of Economic Activities

5.1 The Partial and General Equilibrium Effects of Marriage

We now use the model as a laboratory for counterfactual experiments to shed light on how the marriage incentive affects the spatial distribution of economic activities.

In the first experiment, we eliminate the marriage incentive in two steps. In the first step, we consider the incentives of individuals, holding all general equilibrium objects as given. In the second step, we incorporate the general equilibrium responses.

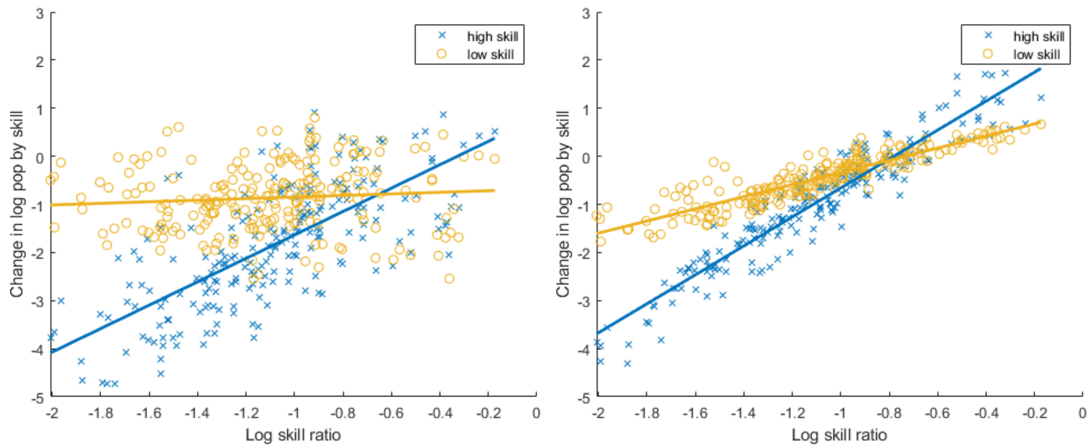
Recall from equation (11) that the expected utility of a city is the sum of the utility if being single and $-\frac{1}{\kappa_M^e}\log(r_{i,M}^{e,\varnothing})$ (for men) or $-\frac{1}{\kappa_F^{e'}}\log(r_{i,F}^{\varnothing,e'})$ (for women). To see where a person would have chosen to live if marriage market outcomes were not part of their consideration, we can simply calculate the expected utility assuming $r_{i,M}^{e,\varnothing} = r_{i,F}^{\varnothing,e'} = 1$. Evaluating individual migration decisions in equation (2) using the expected utility calculated this way gives us the partial equilibrium counterfactual had there been no marriage.

Figure 5a plots the change in the number of people choosing to live in each city from this experiment. The positive slopes of the fitted lines suggest that skill intensive cities are now more attractive to both high- and low-skill workers. The intuition can be found in equation (11)—as skill intensive cities tend to have a higher single rate, removing the marriage incentive makes these cities comparatively more attractive. The slope is less steep for low-skill people primarily because they have higher migration costs and are on average less likely to be born near skill intensive cities than high-skill people. Thus, they are not in an as good position to respond to the change through migration.[22]

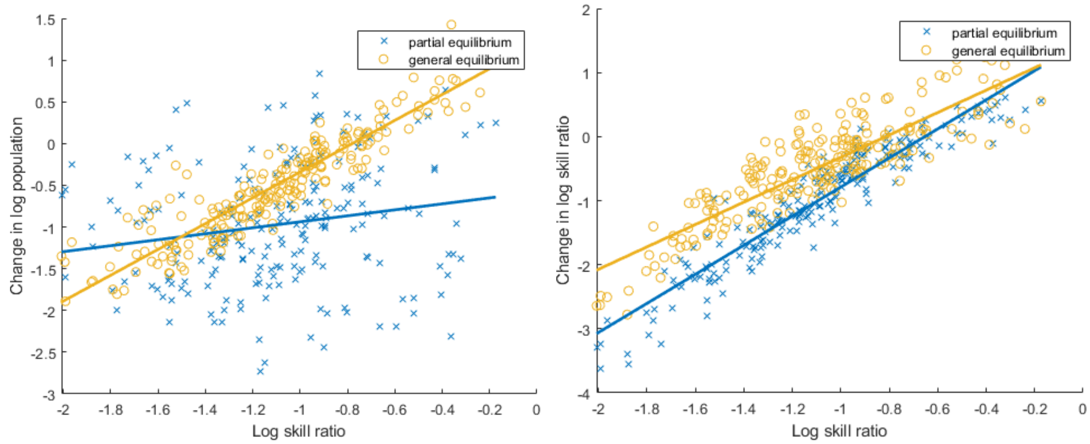
We contrast this partial equilibrium counterfactual exercise with a general equilibrium counterfactual exercise, in which we eliminate marriages by decreasing $\delta_{e,e'}$ to a sufficiently negative number so that virtually no one chooses to be married. Figure 5b plots the results. Two observations are notable. First, compared to Figure 5a, the general equilibrium effect amplifies the reallocation of both types of workers. Second, likely due to the asymmetry in the spillover parameters, the amplification effect is more potent for low-skill workers.

Figure 5c examines how the population of a city changes in the two experiments. As alluded to in the previous discussion, both partial and general equilibrium experiments find the elimination of marriages increases the concentration of the population in the most skill intensive cities, but the strength of the general equilibrium effect, as measured by the slope of the fitted line, is

[22] Another explanation is that in the baseline equilibrium, the gradients of the single rate with respect to city skill ratio differ between high- and low-skill people. To isolate the two explanations, we conduct an alternative partial equilibrium experiment, treating the single rate of low-skill people in a city as if it was the same as the single rate of high-skill people in that city. This exercise rules out that the differential responses of high- and low-skill people are due to their different single rates across cities. We still find skill-intensive city become more skill intensive, which suggests that this result is primarily driven by the difference between high- and low-skill people in migration costs and home states. See Appendix Figure C.2 for more details.



(a) Change in Log Population: *Partial* Equilibrium (b) Change in Log Population: *General* Equilibrium



(c) PE versus GE: City Size (d) PE versus GE: Log Skill Ratio

Figure 5: Partial and General Equilibrium Effects of Marriage

Note: The upper panel plots the impact of the partial and general equilibrium experiments on the location choice of high- and low-skill people. The lower panel compares the impacts of these two experiments on log population and log skill ratio ($\frac{\log(\text{\# of high-skill people})}{\log(\text{\# of low-skill people})}$)

four times as large as the partial equilibrium effect (1.55 versus 0.37). Figure 5d illustrates the impact of the marriage incentive on the skill intensity of a city. In this case, the general equilibrium effect dampens the partial equilibrium effect, but it still leads to an increase in the skill intensity of cities that are already skill intensive.

The marriage incentive as a dispersing force might seem surprising at the first glance, especially given the observation that in the data, high-skill couples are concentrated among the most skill intensive cities. If a high-skilled prefers another high-skilled as a partner and if, as shown in Figure 1a, such outcomes are more likely in skill intensive cities,[23] why doesn't the marriage incentive draw more skills into these cities? What this intuition fails to take into account is that, both the marriage surplus and the division of the surplus between spouses are endogenous

[23] Our model matches these patterns closely. See Appendix Figure C.1.

outcomes that depend on the interaction between the local marriage and labor markets.

The mechanism is best understood through examining the utility in marriages. Figure 6 plots the systematic utility from each type of marriage outcomes for high-skill men and women. For both genders, the utility from staying single and from getting married both increases in the skill intensity of a city, likely due to better amenities and higher wages. However, the utility from getting married increases more slowly. This happens for two reasons. First, because of the diminishing marginal utility from amenities and consumption, embedded in the log utility function, the utility advantage of a married household relative to a single household is smaller in more skill intensive cities, where wages tend to be higher and amenities better. Second, for the utility of a high-skill person marrying a low-skill partner, there is an added effect from competition in the marriage market: in cities with a high skill intensity, low-skill partners, being more scarce, are able to extract a higher share of the marriage surplus. High-skill people, in turn, receive a less favorable division of the surplus in marriages with a low-skill partners in such cities.

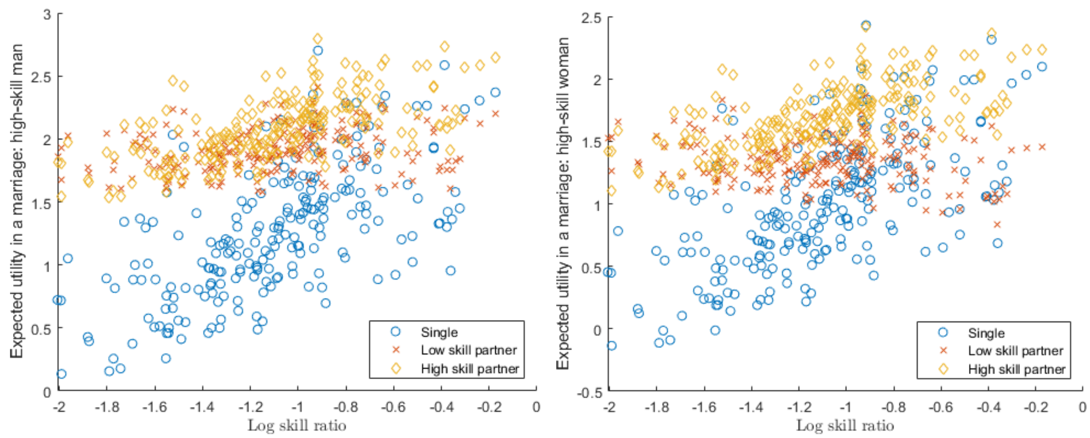
Thus, even though the chance of marrying a high-skill spouse is higher in skill intensive cities and marrying a high-skill spouse is on average more attractive than either staying single or marrying a low-skill spouse, the utility premium from doing so is lower in skill intensive cities. Exactly for this reason, the presence of the marriage incentive makes less skill intensive cities comparatively more attractive. This underscores the importance of incorporating endogenous payoffs from marriages.[24]

To summarize, the first set of experiments find that marriage is a force of dispersion, which drives skills and population to be more evenly spread out. Moreover, the general equilibrium effects through local labor and marriage markets play an important role in modulating this force. In the next subsection, we turn to specific components of the marriage institution and explore the importance of each.

5.2 Decomposing the Marriage Institution

In addition to pooling income from dual-career spouses, the marriage institution in our model manifests itself in three ways. First, it gives couples an option of having a stay-at-home spouse for home production. Second, the value of amenities to a married household is the average of their values to the two spouses, so mixed-skilled households tend to have a more skill-neutral appreciation of amenities. Finally, the two non-economic components, the systematic component $\delta_{e,e'}$ and the idiosyncratic preference draw, also enter household utility. To examine the role of these aspects of the marriage institution, we now shut down these three components separately. Outcomes from these experiments are reported in Table 6.

[24]Ultimately, equation (11) suggests that cities with a higher proportion of people staying single tend to enjoy a lower premium from the marriage market. By combining standard elements from the spatial equilibrium literature (i.e. amenities and housing prices) with a workhorse model for marriage markets, we provide one micro foundation for why this could happen. To the extent that other models of marriage markets can match the spatial distribution of marriage and, in particular, the skill gradient of single rates—as we document empirically—such model would also imply that marriage is a dispersion force.



(a) High-skill Men (b) High-skill Women

Figure 6: Utility from Different Marriage Market Outcomes

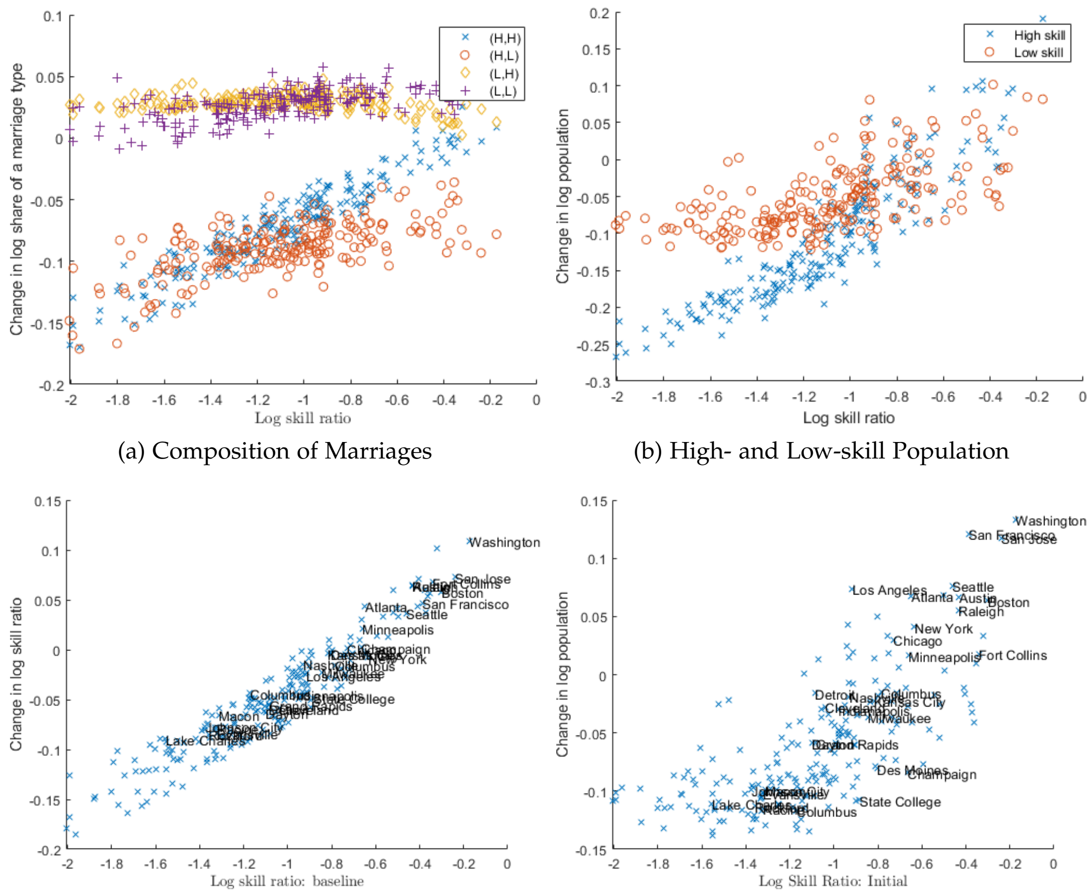
Note: The left panel plots $U^{H,e'}_{d,M}$, for $e' \in \{H, L, \varnothing\}$ (see equation (4)); the right panel plots $U^{e,H}_{d,F}$, for $e \in \{H, L, \varnothing\}$.

Table 6: Changing Specific Components of Marriages

	Baseline	Eliminate the role of				No Marriage
		home production	amenity mix	taste/love	all	
	(1)	(2)	(3)	(4)	(5)	(6)
% married	0.68	0.60	0.68	0.44	0.14	-
the composition of marriages						
(H,H)	0.21	0.20	0.22	0.08	0.05	-
(H,L)	0.13	0.12	0.12	0.15	0.06	-
(L,H)	0.09	0.10	0.09	0.36	0.26	-
(L,L)	0.56	0.58	0.57	0.42	0.62	-
The gradient of $\Delta \log(pop)$ w.r.t. $\log(\frac{H}{L})$	-	0.14	0.06	0.09	0.78	1.55
The gradient of $\Delta \log(\frac{H}{L})$ w.r.t. $\log(\frac{H}{L})$	-	0.10	0.09	0.15	1.10	1.75

Note: The upper panel reports marriage outcomes under different scenarios. Column (1) corresponds to the baseline equilibrium; the remaining columns correspond to different counterfactual equilibria. The lower panel reports the change in the distribution of skills and people across cities between the baseline and each counterfactual scenario, summarized by two gradient metrics.

The option of home production. In the first experiment, we eliminate the option value of home production for married couples by setting $\bar{n}^H$ and $\bar{n}^L$ to sufficiently small numbers. Without this option, marriages become less attractive for all types. The share of married people in population decreases from 68% in the baseline equilibrium to about 60%. The composition of marriages shifts towards low-skill men: both (L, H) and (L, L) types of marriages increase, at the expense of the other two types. These might be moderate changes, but they are accompanied by drastic shifts in the composition of marriages across cities. Figure 7a plots the change in the four types of marriages in each city against the skill intensity of these cities in the baseline equilibrium: whereas the increase in the shares of (L, H) and (L, L) types are modest and similar across cities, the decrease in the shares of (H, H) and (H, L) marriages is much more pronounced and concentrated in cities with a low skill intensity.



(a) Composition of Marriages

(b) High- and Low-skill Population

(c) Skill Ratio

(d) City Size

Figure 7: Changes after Home Production is Eliminated

Note: These figures plot the change in various outcomes when the option of home production is eliminated.

Such disparities arise from the compositional change in local population. Although the reduction in the marriage value hits all cities, it affects the incentive to enter cities with a low skill intensity more heavily because a bigger part of the utility from these cities is due to the marriage incentive. As a result, people gravitate toward more skill intensive cities. Figure 7b plots the change in the number of high- and low-skill workers against the skill intensity of a city in the baseline equilibrium. Both types of people move toward more skill intensive cities, but the response is stronger among high-skill people. In turn, skill intensive cities are getting bigger and more skill intensive, as shown in Figures 7c and 7d.

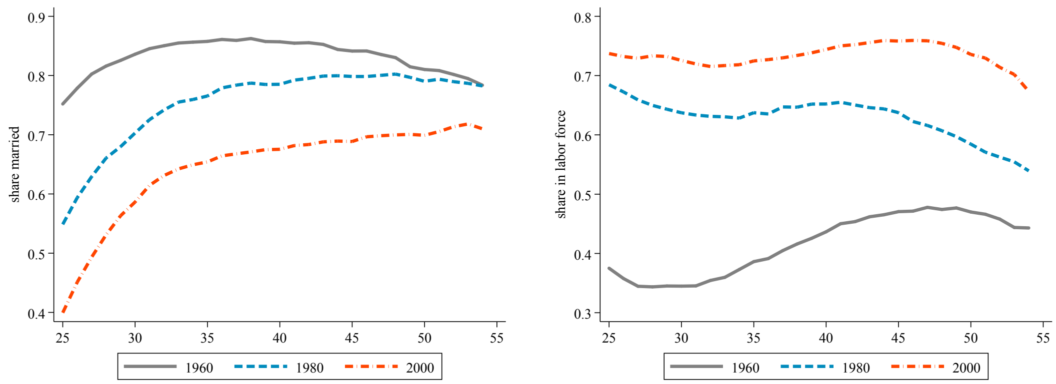
We capture the impact of home production on the spatial concentration of economic activities by two metrics, $\Delta \log(pop)/\log(\frac{H}{L})$ and $\Delta \log(\frac{H}{L})/\log(\frac{H}{L})$, representing the gradients of the change in log population and the change in log skill intensity with respect to the log skill intensity of the baseline equilibrium, which we report in the lower panel of Table 6. These two gradients are 0.14 and 0.1, respectively, accounting for about 9% and 6% of the changes in these two metrics from the complete elimination of marriage (1.55 and 1.75, respectively).

The role of household amenities. Our assumption on household amenities implies the following: for high-skill people residing in cities with amenities appealing to low-skill people—usually the cities with a low skill share—marrying a low-skill people can increase the household's average appreciation of local amenities; the opposite is true for the high-skill people residing in cities with amenities appealing to high-skill people, often the skill intensive cities. The same forces are at play for low-skill people. As such, the mixed-amenities assumption makes marrying the opposite skill type more attractive for individuals living in a city where their own skill type is relatively scarce.

To see how the household amenities assumption affect marriage and location choices, we now assume mixed-skill household enjoy the lower amenities of the two spouses, i.e., $A_d^{e,e'} = \min\{A_d^{e,\varnothing}, A_d^{\varnothing,e'}\}$. The third column of Table 6 shows that this change has little impact on the overall marriage composition. Nevertheless, the spatial distribution of marriages and the skills are affected. As the lower panel of the table reports, measured by the gradient metrics, the change increases the concentration of both population and skills in skill intensive cities, so the preference for household amenities in mixed-skill households is also a dispersion force.

The role of tastes for marriage. In the third exercise, we eliminate the marriage incentive arising from non-economic taste components through two changes. First, we increase κ_s^e to 10, for each $s \in \{M, F\}$, $e \in \{H, L\}$. This weakens the incentive to be in a marriage due to the idiosyncratic taste shocks—-individual choices are now governed by mostly the systematic utility from being either single or in a specific types of marriage.[25] Second, we set $\delta^{e,e'} = 0$, for each $e, e' \in \{H, L\}$, which removes the systematic advantage to be in (H, H), (L, L), or (H, L) marriages and—as the calibrated $\delta^{L,H}$ is negative—the systematic *disadvantage* to be in (L, H) marriages. This change decreases the fraction of people in a marriage to 44%, accompanied by a compositional shift in marriage types toward mixed-skill marriages. Such a composition change tends to make the spatial distribution of skill more even, but its impact is dominated by the overall decrease in marriages. The net result is an increase in the concentration of people in skill intensive cities, as reported in column 4 of Table 6.

All together. Lastly, we combine all three changes in one experiment. As the fifth column reports, the marriage rate now decreases to only 14%. This decrease is larger than the sum of the decrease in the three separate experiments, suggesting the three forces are substitutable motives for marriages.[26] Measuring the impacts of these forces using the two gradient metrics, the three forces together account for around half of the increase in the spatial concentration from a complete removal of marriage.



(a) Share of Married

(b) Female Labor Force Participation

Figure 8: The Changing Marriage Institution: 1960-2000

Note: Panel a shows the share of individuals between ages 25 and 54 who are currently married and not separated. Panel b shows the share of women between ages 25 and 54 who are currently in the labor force.

6 The Changing Marriage Institution and the U.S. Spatial Economy: 1960 to 2000

Like that in many advanced countries, the marriage institution in the U.S. has changed drastically in the second half of the 20th century. As Figure 8a shows, between 1960 and 2000, the share of people who are married declined across all age groups, resulting in a decrease in the overall marriage rate from 83% to 68%.[27] Accompanying this decline was the increasing participation of women in the labor force, which reduces the option value of home production in marriages. Figure 8b shows a cross-the-board increase across all ages in labor force participation of women. This increase is concentrated among married women, as is shown in Appendix Figure A.4.

Over this period, the U.S. economy has also gone through a process of spatial divergence. The share of the 50 most skill intensive MSAs in nationwide population between ages 25 and 54 increased from 27% to 31%; their share of nationwide college graduates increased from 37.5% to 40.5%. This increasing concentration of college graduates is remarkable, considering that there had been a great equalization in access to higher education across space during this period. In light of our earlier finding that marriage is a dispersion force, could the dwindling marriage institution be a reason why the U.S. cities have seen increasing disparities?

[25]Setting $\kappa_s^e = \infty$ would completely eliminate the incentive due to idiosyncratic shocks. But solving the model numerically with $\kappa_s^e \to \infty$ is challenging.

[26]That the joint experiment leads to a bigger decrease than the sum of three separate experiments suggest the marginal effect of individual channels is larger when other channels are not present. In this sense, different channels are substitutes.

[27]This share is calculated among those between 25 and 54 years old. There has been an increase in cohabitation in this age group. We focus on marriage, not cohabitation, because the latter is not consistently defined in the data, and because cohabitation and marriage have different legal definitions, hence not perfect substitutes. Nevertheless, as the increase in cohabitation is small compared with the decline in marriage, even if cohabitation is counted as marriage, the overall marriage rate still declined significantly during the period (see Appendix Figure A.5).

Table 7: Fundamental Changes between 1960 and 2000

	2000		1960	
	Target	Parameter	Target	Parameter
Demographics	15% (M,H), 34% (M,L), 14% (F,H), 36% (F,L)	-	5.7% (M,H),43% (M,L), 3.5% (F,H), 48% (F,L)	-
Marriage patterns	68% people married: HH (21%), HL (13%), LH(9%), LL (56%)	$\delta^{H,H} = 1.06$, $\delta^{H,L} = 0.28$, $\delta^{L,H} = -1.96$, $\delta^{L,L} = 1.20$	83% people married: HH (4%), HL (8%), LH(3%), LL (85%)	$\delta^{H,H} = 1.24$, $\delta^{H,L} = 1.72$, $\delta^{L,H} = -2.13$, $\delta^{L,L} = 3.98$
Gender wage gap	24% for H, 26% for L	$\beta_F^H = 0.76$, $\beta_F^L = 0.74$	36% for H, 38% for L	$\beta_F^H = 0.64$, $\beta_F^L = 0.62$
Female labor force participation	83% among H, 73% among L	$\bar{n}^H = 0.004$, $\bar{n}^L = 1.03$	58% among H, 46% among L	$\bar{n}^H = 0.62$, $\bar{n}^L = 2.19$

Note: This table compares the calibration targets and parameters in 2000 (left) and 1960 (right).

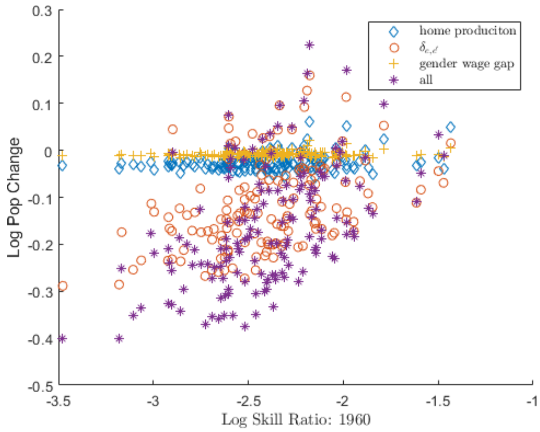
6.1 Calibration to the 1960 Economy

To answer this question, we decompose the changes in fundamentals between 1960 and 2000 and examine their impacts on marriage and the spatial distribution of economic activities.

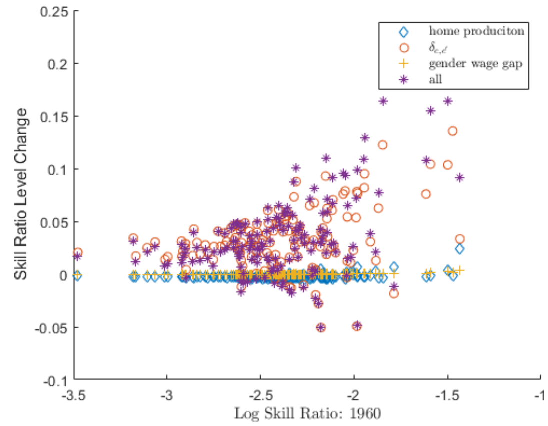
We calibrate the model to match the features of the 1960 economy. Our calibration takes into account that the 1960 economy differs from the 2000 economy in the following aspects: the share of college graduates among men and women, the spatial distribution of the skilled, the composition of marriages, the gender wage gaps, and the female labor force participation rates. Specifically, we feed in the model the composition of people by gender-and-skill in 1960 and the value of p_{n} for 1960, which captures the price and availability of market alternatives to home goods. Between 1960 and 2000, more goods that were primarily produced at home became available on the market. For example, the share of U.S. non-farm employment in the 'eating and drinking places' industry doubled from 3% to 6.7%, reflecting a shift from food at home to food away from home. To capture the lower availability of such options in 1960, we set $p_n = 2$ for 1960.[28] We set $\beta_F^H = 0.64$ and $\beta_F^L = 0.62$ to match the average gender wage gaps in 1960. The remaining parameters are chosen jointly so the model matches the 1960 data exactly, with the sources of identification the same as described in Panel C of Table 5.

Table 7 summarizes the main differences between 1960 and 2000 and the resulting change in the calibrated parameters. Aside from the share of married people, the composition of marriages also changed over this period. In particular, in 1960, there were more (L, L) and fewer (H, H) marriages. Through the lens of the model, such compositional changes cannot be entirely accounted for by the increase in college shares alone—the calibrated $\delta^{L,L}$ for 2000 is lower than that for 1960, suggesting that the shift in marriage preferences also plays a role in the decreasing share of (L, L) marriages. Compared with their counterparts in 2000, $\delta^{H,L}$ is larger in 1960 but

[28] A more liberal interpretation of the decrease in p_{n} is the increasing quality and decreasing price of home appliances, which reduces the time required for housework. During this period, the price of major home appliances decreased by more than 50% (Greenwood et al., 2005). Ultimately, because p_{n} and the $\bar{n}^e$ jointly determine the labor force participation, different values of p_{n} for 1960 will affect the calibrated $\bar{n}^e$ for 1960, but conditional on that, not other implications of the model. We show in Appendix Table C.2 that the model's predictions on spatial divergence are insensitive to the value of p_{n} for 1960.

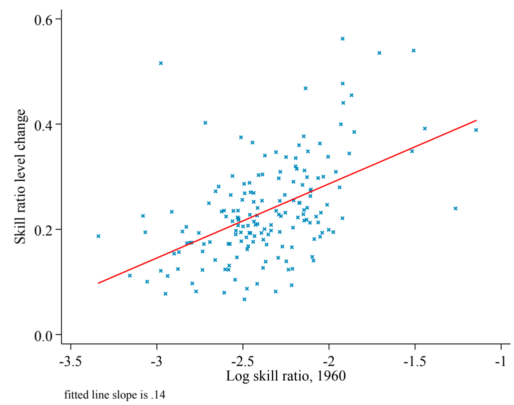


(a) Change in City Size, 1960-2000: Model



(b) Change in Skill Intensity, 1960-2000: Model

(c) Change in City Size, 1960-2000: Data



(d) Change in Skill Intensity, 1960-2000: Data

Figure 9: Change between 1960 and 2000: Data versus Model

Note: The upper panel plots the change in endogenous outcomes after we set some parameters in the 1960 economy to their values in the 2000 economy. The lower panel plots the actual change between 1960 and 2000 in population and skill share by initial city skill ratio. The slope of the fitted line is 0.33 in Figure 9c and 0.14 in Figure 9d.

$\delta^{L,H}$ is smaller. This could reflect the shift away from the stereotype that the husband should always be the breadwinner of the family.

In the model, the increase in female labor force participation could have happened for three reasons. First, a decrease in the gender wage gap encourages more women to participate in the labor market, which we capture in β^e_F. Second, the increase in the market provision of home goods leads to the halving of p_n, the value of which is assigned externally. Third, the wedges for home production, introduced via $\bar{n}^H$ and $\bar{n}^L$, which are calibrated to match the residual change in female labor force participation. It turns out that the first two forces are not enough to generate the increasing labor force participation, so our model infers $\bar{n}^H$ and $\bar{n}^L$ have decreased significantly over the four decades.

Table 8: Accounting for Changes between 1960 and 2000

	Model				Data
	(1) home production ($p_n, \bar{n}^e$)	(2) gender wage gap (β_F^e)	(3) $\delta^{e,e'}$	(4) all together	(5)
The gradient of skill	0.003	0.001	0.028	0.054	0.14
The gradient of population	0.015	0.003	0.11	0.20	0.33

Note: The table reports the changes in the population and skill distribution summarized by the two gradient metrics in four counterfactual scenarios, and the actual change in these metrics between 1960 and 2000.

6.2 Counterfactual Experiments: 1960 to 2000

We examine the extent to which the changes in the calibrated fundamental parameters affect the economy. On the basis of the 1960 economy, we apply three changes, first separately and then jointly. In the first experiment, we set the two parameters governing home production, p_{n} and $\bar{n}^e$ to the values in the 2000 baseline economy. The second and third experiments set the gender wage gaps β_F^e and the non-economic marriage returns $\delta^{e,e'}$ to the 2000 values, respectively. Finally, we combine all these changes at once.

Figures 9a and 9b plot the results. The changes in home production and the gender wage gap, on their own, have relatively small impacts. On the other hand, the change in $\delta^{e,e'}$ leads to a reshuffling of population, especially of high-kill people, towards skill intensive cities. As a result, notwithstanding a few exceptions, skill intensive cities see an increase in both size and skill intensity. When the change in $\delta^{e,e'}$ is combined with the change in other fundamentals, the patterns become slightly stronger.

This increasing concentration of population and skills is qualitatively in line with the data. Figures 9c and 9d show that over this period, on average, skill intensive cities experienced more rapid growth in both the overall population and the level of its skill shares.

We use the gradients of the changes in log population and in skill ratio with respect to the initial log skill ratio to summarize how the changes in fundamentals between 1960 and 2000 affect the spatial economy. Table 8 summarizes the results. The change in home production and the gender wage gap, on their own, have little impact, consistent with what the figures show. But when combined with the change in non-economic marriage returns, they generate a stronger joint effect. Overall, the three forces related to the marriage market can account for one-third of the skill divergence and half of the increasing population concentration in skill intensive cities.

In an influential paper, Costa and Kahn (2000) hypothesize that the increase in the share of dual-career households is an important driver of the increasing concentration of power couples in big cities. As big cities also tend to be skill intensive, this seems to suggest that the rising female labor force participation should *in itself* have a significant impact on spatial divergence, contrary to what we find in Figure 9a and Table 8.

The empirical analysis in Costa and Kahn (2000) is to estimate location choice for different types of *households*, with the key implicit assumption that individuals *first* marry and decide whether the wife works, and *then* make the migration decision. The finding that dual-career power couples have a higher and increasing probability of residing in large cities than comparison groups lead to the conclusion that the increase in dual-career households explains the concentration of power couples in big cities. While this channel is certainly at play, note that in the data, people over the initial marriage age have a low migration rate, so the migration decision of dual-career couples might not be the dominant force that shapes the size of cities. A more likely explanation for the over-representation of power couple in big cities is that young, educated men and women are more likely to move to those cities as singles, which increases their probability of matching with a skilled partner.[29]

Instead of modeling the location choice for married households, we focus on the decision of young adults, who account for most of the inter-MSA migration. In this model, the impact of the changing social norm about working wives on location choices is summarized by the female labor force participation rate. Specifically, by combining equations (14) and (15), we have

$$\overline{V}_d^{e,e'} = \underbrace{\delta^{e,e'} + \log(A_d^{e,e'}) - \alpha \log(r_d) + \frac{\bar{\gamma}}{\eta_F^{e'}} + \log\left(\exp(c_1) \cdot I_{d,W}^{e,e'} \cdot p_h^{-\beta}\right)}_{\text{household utility conditional on the wife working}} - \frac{1}{\eta_F^{e'}} \log(l_d^{e,e'}),$$

i.e., the household utility in city d equals the utility when the wife chooses to work, adjusted by $-\frac{1}{\eta_F^{e'}}$ times the log of the participation rate. Because in our model (and in the data, see Appendix Figure A.6), the increase in female labor force participation over this period is similar across cities, this force by itself affects the expected household utility of cities in a similar way without having a first-order impact on spatial divergence.

7 Concluding Remarks

In this paper we have presented a quantitative spatial model with endogenous marriage formation, in which local labor markets, local marriage markets, and local spillovers interact to determine individuals' location and marriage decisions. Marriage in the model embodies both a non-economic component, capturing the 'love' between the couple, and an economic component, arising from household consumption and labor supply decisions. We show that the model can account for the spatial heterogeneity in the U.S. local marriage markets.

We have used the model to study the role of the marriage institution and its changes over the second half of the 20th century in shaping the spatial distribution of economic activities. One salient finding of our paper is that marriage is a force of dispersion that reduces the concentration of population and skills in skill intensive cities. The heterogeneous and endogenous returns from the local marriage market is the key channel driving this result. We also show that the declining importance of marriage since the 1960s can account for a third to a half of the observed divergence among cities.

[29] Compton and Pollak (2007) make a similar point. Using individual panel data and the 2000 census, which was not available to Costa and Kahn (2000), they find that the concentration of power couples in major cities is better explained by larger probabilities of power couple formation in these cities than by the migration of married power couples to these cities.

In the empirical application, we have limited our attention to the U.S. But the decline in the marriage rate was common to most industrialized countries. Our model provides a tractable tool for analyzing how such trends affect the economic geography in those countries. The sufficient statistic we derive for the marriage market premium of cities can be of use in measuring and comparing the importance of marriages for location choices across countries and over time.

To highlight the central tradeoff associated with marriages in the spatial equilibrium setting, we have kept the model simple. In particular, the model being static, it abstracts away from the possibility of divorces and life-cycle decisions of migration. Since the main questions we seek to answer are about long-run outcomes, such simplifications, we believe, allow us to focus on key interactions between marriage and labor markets without losing important insights. A growing strand of the literature has examined the dynamic implications of spatial equilibrium models (e.g., Caliendo et al., 2019). Future work should incorporate dynamic marriage decisions into such models to shed light on how location and marriage considerations interact over the life cycle and what implications such interactions have for cities.

References

Allen, Treb and Costas Arkolakis, "The Welfare eEfects of Transportation Infrastructure Improvements," *NBER Working Paper No. 25487*, 2019.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 2010, *100* (3), 1008–45.

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

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

Blau, Francine D and Lawrence M Kahn, "Changes in the Labor Supply Behavior of Married Women: 1980–2000," *Journal of Labor Economics*, 2007, *25* (3), 393–438.

Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-market Impact of Immigration: Theory and Evidence From the United States," *Econometrica*, 2020, *88* (3), 1071–1112.

Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte, "The Impact of Regional and Sectoral Productivity Changes on the US Economy," *The Review of Economic Studies*, 2014.

_ , Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 2019, *87* (3), 741–835.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss, "Partner Choice, Investment in Children, and the Marital College Premium," *American Economic Review*, 2017, *107* (8), 2109–67.

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

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

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

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, "Income Growth and the Distributional Effects of Urban Spatial Sorting," *NBER Working Paper 26142*, 2019.

Davis, Donald R and Jonathan I Dingel, "A Spatial Knowledge Economy," *American Economic Review*, 2019, *109* (1), 153–70.

Decker, Colin, Elliott H Lieb, Robert J McCann, and Benjamin K Stephens, "Unique equilibria and substitution effects in a stochastic model of the marriage market," *Journal of Economic Theory*, 2013, *148* (2), 778–792.

Diamond, Rebecca, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, March 2016, *106* (3), 479–524.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 2018, *108* (4-5), 899–934.

Eckert, Fabian, "Growing apart: Tradable Services and the Fragmentation of the US Economy," *mimeograph, Yale University*, 2019.

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

Fajgelbaum, Pablo D and Cecile Gaubert, "Optimal Spatial Policies, Geography, and Sorting," *The Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.

_ and Edouard Schaal, "Optimal Transport Networks in Spatial Equilibrium," *NBER Working Paper No. 23200*, 2019.

Galichon, Alfred and Bernard Salanié, "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *Available at SSRN 1804623*, 2020.

Gaubert, Cecile, "Firm Sorting and Agglomeration," *American Economic Review*, 2018, *108* (11), 3117–53.

Gemici, Ahu, "Family migration and labor market outcomes," *Mimeo*, 2011.

Giannone, Elisa, "Skilled-biased Technical Change and Regional Convergence," *University of Chicago. Unpublished manuscript*, 2017.

Graham, Bryan S, "Comparative Static and Computational Methods for an Empirical One-to-one Transferable Utility Matching Model," in "Structural Econometric Models," Emerald Group Publishing Limited, 2013.

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

_ , Nezih Guner, and Guillaume Vandenbroucke, "Family Economics Writ Large," *Journal of Economic Literature*, 2017, *55* (4), 1346–1434.

Hsieh, Chang-Tai and Enrico Moretti, "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 1–39.

Jiao, Yang and Lin Tian, "Geographic Fragmentation in a Knowledge Economy," 2019.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, migration, and local employment elasticities," *American Economic Review*, 2018, *108* (12), 3855–90.

Nelson, Jon P, "Residential Choice, Hedonic Prices, and the Demand for Urban Air Quality," *Journal of Urban Economics*, 1978, *5* (3), 357–369.

Redding, Stephen J., "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 2016, *101*, 148–167.

Redding, Stephen J, "Suburbanization in the United States 1970-2010," *NBER Working Paper No. 28841*, 2021.

_ and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9*, 21–58.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 1974, *82* (1), 34–55.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman, "Cognitive Hubs and Spatial Redistribution," *National Bureau of Economic Research*, 2019.

Ruggles, Steven, J Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B Schroeder, Matthew Sobek et al., "Integrated Public Use Microdata Series," *Minneapolis: University of Minnesota*, 2021, *42.*

Saiz, Albert, "The Geographic Determinants of Housing Supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Venator, Joanna, "Dual-Earner Migration Decisions, Earnings, and Unemployment Insurance," *Washington Center for Equitable Growth Working Paper*, 2021.

ONLINE APPENDIX FOR
The Dual Local Markets: Family, Jobs, and the Spatial Distribution of Skills

Jingting Fan[*] Ben Zou[†]

This version: November 2021

Contents

[*]Pennsylvania State University; jxf524@psu.edu.
[†]Michigan State University; benzou@msu.edu.

A Data and Additional Empirical Results

A.1 Data and Sample

Data sources. The main data used in this paper are the 5% samples of the 1960 and 2000 population censuses from the IPUMS (Ruggles et al., 2021), which are based on the long form of population censuses. They are nationally representative samples and include detailed geographic, demographic, and economic characteristics at the individual and household levels.

Definition of local markets. For quantitative analysis, we define local marriage and labor markets as metropolitan statistical areas (MSAs), which have long been used as the definition of local labor markets. There are 283 identifiable MSAs in the 2000 census and 177 in the 1960 census. The delineations of MSAs change over time. When comparisons of local markets across years are needed, we focus on MSAs that are consistently defined across the two censuses.

An alternative definition of a local market is a commuting zone. Commuting zones cover the entirety of the United States. For main quantitative analyses, we use MSAs instead of commuting zones primarily because a key set of parameters in our model—local land supply elasticities, which is pinned down by indices constructed from land area, geographic features, and zoning regulations—are only available for MSAs (Saiz, 2010; Diamond, 2016).

All reduced-form patterns reported in the paper and this appendix are robust if we define local markets using commuting zones.

Definitions of skill and marriage. A person is high skilled if they appear in Census 2000 and have a college degree; or they appear in Census 1960 and have 4 or more years of college education. Marital statuses classified by the IPUMS include married with or without spouse present, separated, divorced, widowed, or never married. We call an individual married if they report that they are currently married, with or without spouse present.[1] Finally, in showing the change in the marriage rate between 1960 and 2000, we also consider an alternative definition of marriage, which includes cohabitation (see Figure A.5).

In tabulating the composition of households by type, we need to classify couples by spouses' skill levels. For this purpose, we need to identify who is married to whom within a household. For accurate measurements, we focus only on the head couple of a household.[2] We also only focus on traditional families where a marital couple consists of a man and a woman.

Choices of sample groups for estimation. In reality, people make migration, marriage and labor market decisions throughout their lifetime and not in a strict order. For simplicity, we have developed a static model, in which individuals live through three life stages: a pre-period

[1] This definition implicitly treats divorcees and those who are separated from their spouses as singles, which is consistent with the interpretation of our model as for 'successful marriages.' We note that the patterns of single rates and the composition of marriages across cities are similar if we instead focus only on the never-married.

[2] There may be multiple married couples in a household, but often we are unable to identify both spouses. For example, a married son may live with his wife and his parents (the head couple). He is identified as a child of the head, and we can identify that he is a son based on his gender. His wife will be classified as the daughter-in-law of the head. But we cannot identify the couple if there are other sons in the household, so we omit such cases. Such omission is likely negligible because in cases like this, the son and his wife is usually identified as a separate household from the parents, even if they live in the same house. Therefore, co-existence of multiple married couples in a same household is rare.

when individuals choose which city to live in, a period when individuals engage in the marriage market, and one that is after the marriage market is settled and families are formed. Such assumption fits the empirical regularities that the migration rate is highest in one's early 20s and declines rapidly with age, especially after one gets married and has children; similarly, the probability of entering marriage is highest between one's late 20s and early 30s.

In quantification, we estimate parameters governing decisions in different life stages by restring the sample to different age groups that are consistent with the model. In particular, we estimate the female labor force participation (Section 4.2) and marriage preferences (Section 4.3) using people between ages 40 and 54, a group that is above the typical ages of initial marriages but before retirement. When estimating migration costs (Section 4.4), we focus on people between ages 25 and 39.

A.2 Empirical Patterns: Robustness and Additional Results

Marriage rate and spatial divergence in OECD countries. One observation of the paper is that the declining marriage rate goes hand in hand with the increasing spatial concentration in economic activities. Our paper focuses on the United States, but this pattern is prevalent across industrialized countries.

Table A.1 documents relevant statistics for the OECD countries. Between 1970 and 2010, all countries with available data experienced a decrease in the net marriage rate.[3] Accompanying this trend was the increasing concentration of population in big cities, measured using the share of nationwide population residing in large metropolitan areas. We define large metropolitan areas as those among the top 40% of a country's metropolitan size distribution. We obtain countries' metropolitan size over time from the United Nation dataset on the size of urban agglomerations, which includes all urban areas with above 300,000 population in 2018. Table A.1 shows that 26 out of 35 OECD countries saw an increase in the share of their major cities. Our model has the potential in explaining the broad pattern observed in many high-income countries.

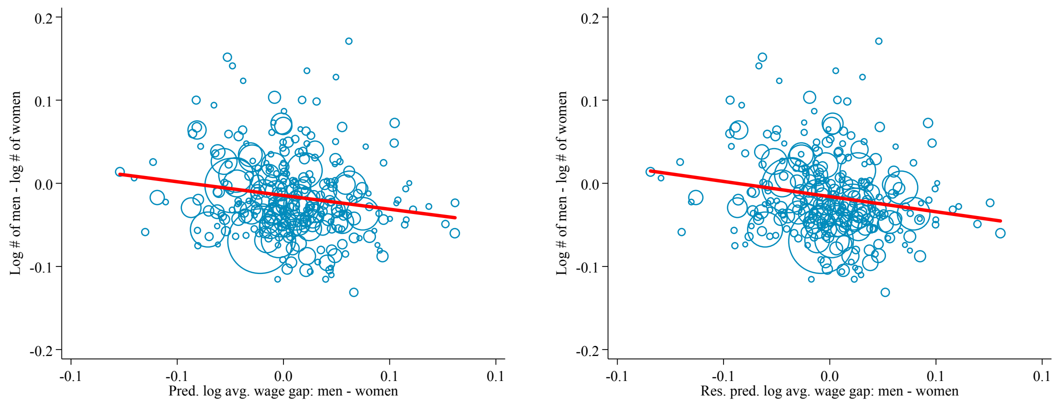
Table A.1: Marriage Rate and Spatial Divergence: OECD countries

Country	Net marriage rate (# per 1,000 people) Δ btw.1970-2010	Share of pop. in large metro areas (%) Δ btw.1970-2010	Country	Net marriage rate (# per 1,000 people) Δ btw.1970-2010	Share of pop in large metro areas (%) Δ btw.1970-2010
Australia	-5.20	-3.40	Japan	-5.60	12.01
Austria	-3.30	-1.01	Latvia	-3.60	0.85
Belgium	-5.70	1.70	Lithuania	-4.50	5.29
Canada	-4.60[1]	7.12	Mexico	-2.20	10.43
Chile	-4.60[2]	9.60	Netherlands	-6.20	-2.08
Colombia	-	13.39	New Zealand	-5.20	7.23
Costa Rica	-1.0[2]	10.84	Norway	-4.00	1.78
Czech Republic	-5.40	0.76	Poland	-3.10	0.28
Denmark	-2.50	-6.50	Portugal	-8.10	5.66
Estonia	-4.30	2.24	Korea	-4.60	17.27
Finland	-4.40	9.99	Slovakia	-4.60	1.40
France	-0.70[3]	0.63	Spain	-3.70[2]	1.54
Germany	-3.70	-0.36	Sweden	-1.00	0.48
Greece	-3.40	0.24	Switzerland	-3.90	5.07
Hungary	-5.90	-1.48	Turkey	-	17.84
Ireland	-2.50[2]	-1.99	United Kingdom	-5.10	-3.54
Israel	-3.50	11.53	United States	-3.90	0.24
Italy	-4.20	-0.50			

Note: The sample includes OECD countries. Net marriage rate is the crude marriage rate (number of marriages per 1,000 people) minus the crude divorce rate (number of divorces per 1,000 people). Marriage rate and divorce rate data are from the OECD database, Table SF-3-1. Data from select countries are defined slightly differently and indicated by superscripts: [1]1970-2000; [2]raw marriage rate; [3]1995-2018. Population in large metro areas are from the urban agglomeration dataset maintained by the United Nations (population.un.org/wup/Download). 'Large' is defined as among the top 40% biggest cities of a country in the dataset, which include all urban agglomerations with a population more than 300,000 in 2018. Nationwide population is from the World Bank's World Development Indicators.

[3]Data for Canada, Chile, Ireland and Spain are either over a slightly different period, or defined using different variables. See the table note for details.

The relationship between the predicted gender earnings difference and the gender ratio. We use Figure 1b in the paper, which shows that there is no correlation between the relative earnings between men and women in a labor market and the gender ratio, as suggestive evidence that marriage prospects incentivize people to move to places where their earnings potential is not maximized. One concern with this piece of evidence is that the gender earnings difference can be affected by the relative labor supply. To rule out this possibility, we construct the predicted earnings by gender that are derived from the industry-and-occupation structure in the local market.



(a) Predicted Wage

(b) Pred. Wage, Adj. for City Size and Skill Ratio

Figure A.1: Gender Ratio and Predicted Gender Wage Difference

Note: Each bubble represents an MSA in 2000. We restrict the sample to full-time workers between ages 25 and 54. To predict wage, we regress, separately for workers by gender and skill, log annual earnings on a flexible function of age, education, their interactive terms, and a set of industry-occupation indicators. Then we aggregate industry-occupation fixed effects (using their shares of local employment as weights) separately by gender in each MSA. The predicted gender wage gap in an MSA is the difference between men and women in the weighted average of industry-occupation fixed effects. In Panel a, the slope is -0.34 with a robust standard error of 0.13. In Panel b, log gender wage gap is further residualized from log city population size and skill ratio. The slope is -0.36 with a robust standard error of 0.11.

Specifically, from the 2000 census, we obtain the sample of full-time workers between ages 25 and 54. We regress, separately for workers by gender and skill, log annual earnings on flexible functions of age, education, their interactive terms, and a set of industry-occupation indicators. We then use these industry-occupation fixed effects and the weight of different industry-occupation cells in a city to generate wage predictions for men and women in that city. Figure A.1a plots the difference in predicted wages between men and women against the gender ratio. Since the wages are predicted using the industry-occupation structures of cities, the figure can be viewed as the relatively supply curve. It shows that relative prices and relative supply in the local market are not positively correlated, further supporting the hypothesis that marriage market incentives play an important role in people's decision on where to live.[4] One may still be

[4]If anything, the correlation is negative. A hypothesis that can explain this negative correlation is that women might

concerned that other characteristics of the city, such as its size and skill ratio, drive this relationship. In Figure A.1b, we further residualize the predicted log wage difference by regressing it on log population size and log skill ratio. The pattern is essentially unchanged.

Log skill ratio and the probability of being single. Figure 4f of the paper shows a positive correlation between the single rate and the log skill ratio of local markets. One might be concerned that such a pattern could be driven by differences in age compositions across cities. To address this concern, we use individual level data and regress the indicator for being single on city skill ratio, separately for each gender-skill group and controlling for age-by-race fixed effects. Figure A.2 visualizes the finding. The slope of the fitted lines for all groups are positive, economically meaningful, and statistically significantly above zero. On average, increasing the log skill intensity (H/L) by one is associated with a 5.5 percentage point (p.p.) increase in the probability of staying single.

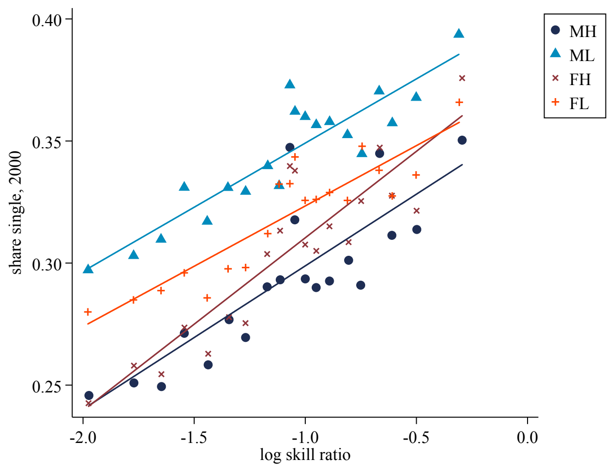


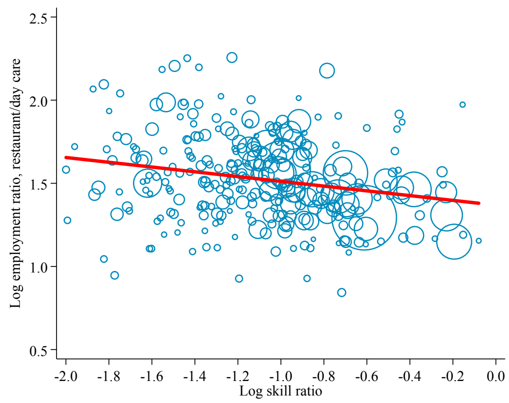
Figure A.2: Prob. of Being Single by Gender-skill, against Local Log Skill Ratio
Note: The sample includes individuals between ages 25 and 54 from Census 2000. The graph shows binned scatter plots and linearly fitted lines separately for each gender-skill group, adjusting for age-by-race fixed effects.

Log skill ratio and differential amenities for singles and married couples. As discussed in the paper, one additional channel for marriage rates to be lower in skill intensive cities is that amenities in these cities appeal more to single people, which can be tractably introduced to the baseline model by assuming $\delta^{e,e'}$ to be a decreasing function of city skill intensity.

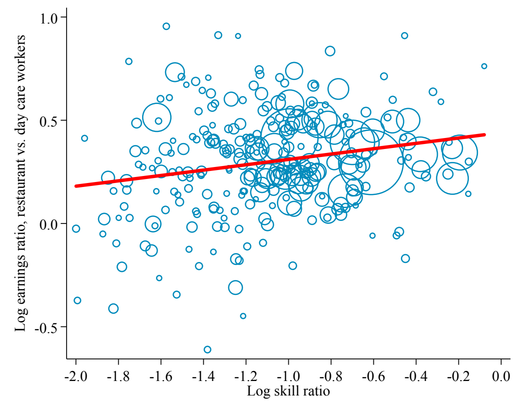
While it is difficult to categorize various amenities to be 'singles-friendly' and 'family-friendly,' it seems reasonable to assume that compared to singles, married couples derive higher utilities fromm the access to affordable childcare than they do from restaurants and bars. We proxy for the accessibility these amenities by the employment in corresponding industries; we proxy for those services using the average earnings of workers in these industries.

be attracted to places where the earnings potential of men is higher for marital reasons (Edlund, 2005). Although this hypothesis is also consistent with our model, we note that the negative slope is relatively weak and our main point is on the lack of a positive slope.

Figure A.3a shows that skill-intensive cities do not necessarily have more restaurants and bars relative to daycare facilities—if any, they have relatively more employment in daycare facilities; Figure A.3b further shows that restaurants are likely more expensive in skill-intensive cities relative to daycare centers. Thus, if we were to use either the accessibility or the cost proxy to estimate the relationship between 'singles-bias' of amenities and log skill ratio of cities, we would have found that this mechanism tends to generate *lower* single rates in skill intensive cities. For this reason, incorporating this mechanism is unlikely to undermine the importance of the main mechanisms in our baseline model.[5]



(a) Differential Amenities: Accessibility



(b) Differential Amenities: Cost

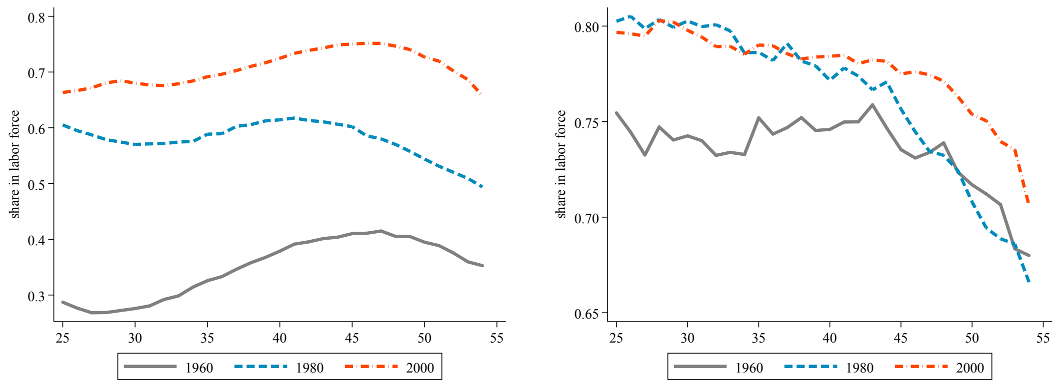
Figure A.3: Amenities Attractive to Singles and Married Couples

Note: Each bubble represents an MSA in 2000. The size of the bubble is proportional to MSA's population. We pick two industries to represent amenities that are particularly valued by singles and married couples. We postulate that unmarried people derive a relatively high value from restaurants and bars (eating and drinking places), while married people derive a relatively high value from childcare services (child daycare centers). Panel (a) shows that more skill-intensive cities do not have amenity compositions that favor singles: number of workers in the eating and drinking industry relative to that of workers in the child daycare industry declines with city's skill intensity. Panel (b) shows that the average earnings in the eating and drinking industry relative to that in the child daycare industry—which serves as a proxy for the relative costs of these two services—increases with the city's skill intensity.

Secular trends in the female labor force participation by marital status. Figure 8b in the paper shows that women's labor force participation has increased substantially between 1960 and 2000. Figure A.4a and Figure A.4b show changes in women's labor force participation between 1960 and 2000 by age and by marital status. Much larger increases are observed among married women, whose labor force participation rates increased between 30 and 40 p.p., depending on the age. Unmarried women's labor force participation also increased, but by a much lesser degree (around 5 p.p.). This observation motivates the focus of our model on the labor force participation decision of married women only.

Incorporating cohabitation in marriages. Figure 8a in the paper shows that the share of

[5] This evidence is suggestive because the quality of services are not accounted for, and we only consider two modal amenities, from which married and single families may derive substantially different utility.



(a) Married (b) Unmarried

Figure A.4: Female Labor Force Participation, 1960-2000, by marital status

Note: Data from the 1960, 1980, and 2000 censuses. The graphs show the share of women between ages 25 and 54, by age and marital status, who are currently in the labor force. Panel a shows those who are currently married. Panel b shows those who are not married.

people in a marriage has been declining for all age groups in the second half of the 20th centruy. During the same time period, there has been an increase in cohabitation. We focus on marriages because cohabitation is less consistently defined in the data. Here we show in Figure A.5 that including cohabitation only slightly offset the reduction in marriages.

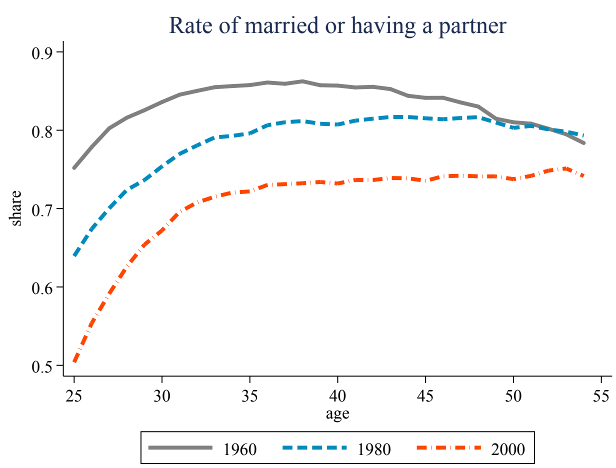
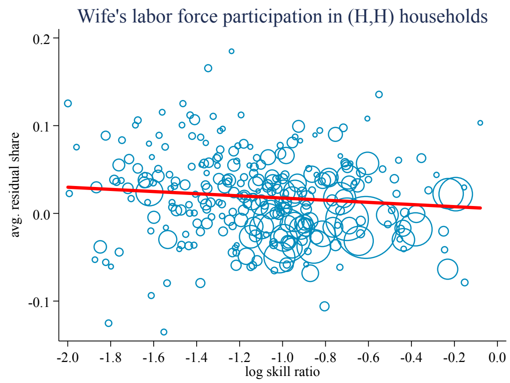


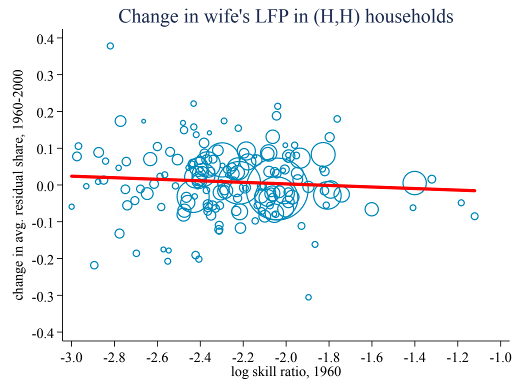
Figure A.5: Share of Married or Cohabiting, 1960 - 2000

Note: The sample includes those between 25 and 54 years old in the 1960, 1980, and 2000 censuses. The figure plots the share of people who is either married or cohabiting. A person is considered married if he or she is married with or without the spouse present. A person is considered as cohabiting with the head of the household if he or she is defined as a 'partner or friend' in 1960, a 'partner or roommate' in 1980, or a 'unmarried partner' in 2000.

Wife's labor supply among power couples, by local skill composition. In Section 6.2 we explain that one reason our finding differs from that of Costa and Kahn (2000) is that, in our model, the effect of increasing female labor force participation on location choice depends on the



(a) 2000



(b) Change between 1960 and 2000

Figure A.6: Share in Labor Force Among Wives in 'Power Couple' Households

Note: Each bubble represents an MSA. The sample includes wives in households in which both the husband and the wife have college degrees. Labor force participation is adjusted by age and the presence and the number of children under 6. Panel a shows the relationship between log skill ratio and wives' labor force participation rate among MSAs in 2000. Panel b shows the relationship between the log skill ratio in 1960 and the change in wives' labor force participation rate among consistently defined MSAs between 1960 and 2000, thus a smaller number of MSAs in panel b.

extent to which the participation rate differs across cities. Contrary to the common perception, however, Figure A.6a shows that the labor force participation of wives in power couples is not higher in more skill intensive cities. Figure A.6b further shows that the *increase* in the wife's labor force participation rate among power couples between 1960 and 2000 is also not faster in initially more skill-intensive cities. Through the lens of our model, this pattern implies the increasing participation by itself does not have a first-order impact on spatial divergence.

B Theory

B.1 Proof of Lemma 1

Proof. Setting $U_{d,M}^{e,\varnothing} = \overline{V}_d^{e,\varnothing}$ and $U_{d,F}^{\varnothing,e'} = \overline{V}_d^{\varnothing,e'}$, one can verify that part 1 of the lemma holds. Below we prove parts 2 and 3 of the lemma.

Consider two couples in the stable match with the exact same skill composition, i.e., (ω, ω') and $(\tilde{\omega}, \tilde{\omega}')$ such that the two husbands ω and $\tilde{\omega}$ has the same skill e, and the two wives $\tilde{\omega}$ and $\tilde{\omega}'$ has the same skill e'. From conditions 4 and 5 of Definition 1 in the text, we have:

$$\begin{aligned}
u_d^{M,e}(\omega) + u_d^{F,e'}(\omega') &= \overline{V}_d^{e,e'} + \xi_M^{e,e'}(\omega) + \xi_F^{e,e'}(\omega) \\
u_d^{M,e}(\omega) + u_d^{F,e'}(\tilde{\omega}') &\geq \overline{V}_d^{e,e'} + \xi_M^{e,e'}(\omega) + \xi_F^{e,e'}(\tilde{\omega}') \\
u_d^{M,e}(\tilde{\omega}) + u_d^{F,e'}(\tilde{\omega}') &= \overline{V}_d^{e,e'} + \xi_M^{e,e'}(\tilde{\omega}) + \xi_F^{e,e'}(\tilde{\omega}') \\
u_d^{M,e}(\tilde{\omega}) + u_d^{F,e'}(\omega') &\geq \overline{V}_d^{e,e'} + \xi_M^{e,e'}(\tilde{\omega}) + \xi_F^{e,e'}(\omega')
\end{aligned} \tag{B.1}$$

Taking the difference between the first two and the last two lines of equation (B.1) gives

$$\xi_F^{e,e'}(\omega) - \xi_F^{e,e'}(\tilde{\omega}') \geq u_d^{F,e'}(\omega') - u_d^{F,e'}(\tilde{\omega}') \geq \xi_F^{e,e'}(\omega) - \xi_F^{e,e'}(\tilde{\omega}'),$$

which implies $u_d^{F,e'}(\omega') - \xi_F^{e,e'}(\omega) = u_d^{F,e'}(\tilde{\omega}') - \xi_F^{e,e'}(\tilde{\omega}')$, that is, the difference between the utility of the wife and their idiosyncratic preference for that assignment is a constant independent of the draws of both the husband and the wife for that assignment. Similarly, the same holds for husbands, i.e., $u_d^{M,e}(\omega) - \xi_M^{e,e'}(\omega) = u_d^{M,e}(\tilde{\omega}) - \xi_M^{e,e'}(\tilde{\omega})$.

Define $U_{d,F}^{e,e'} \equiv u_d^{F,e'}(\omega') - \xi_F^{e,e'}(\omega)$ and $U_{d,M}^{e,e'} \equiv u_d^{M,e}(\omega) - \xi_M^{e,e'}(\omega)$, these two scalers then satisfy conditions 2 and 3 of the lemma. □

B.2 Proof of Proposition 1

Proof. We prove this proposition for a male ω of skill e married to a woman of skill e' in the stable match. Proof for other types of individuals is analogous. By Lemma 1, we have:

$$u_d^{M,e}(\omega) = U_{d,M}^{e,e'} + \xi_M^{e,e'}(\omega).$$

Lemma 1 and condition 2 of Definition 1 together imply:

$$u_d^{M,e}(\omega) \geq \overline{V}_d^{e,\varnothing} + \xi_M^{e,\varnothing}(\omega) = U_{d,M}^{e,\varnothing} + \xi_M^{e,\varnothing}(\omega).$$

It remains to show that

$$U_{d,M}^{e,e'} + \xi_M^{e,e'}(\omega) \geq U_{d,M}^{e,e''} + \xi_M^{e,e''}(\omega),\ e'' \neq e.$$

Suppose the opposite is true, then ω can 'lure' away a woman of type e'' who is in a type (e, e'') type marriage in the stable match, which violates condition 5 of Definition 1. This completes the proof of equation (6). □

B.3 Deriving Equation (28)

From equations (8) and (9), we have:

$$\frac{q_d^{e,e'}}{q_d^{e,\varnothing}} = \frac{r_{d,M}^{e,e'}}{r_{d,M}^{e,\varnothing}} = \frac{\exp(\kappa_M^e \cdot U_{d,M}^{e,e'})}{\exp(\kappa_M^e \cdot U_{d,M}^{e,\varnothing})} \Rightarrow q_d^{e,e'} = q_d^{e,\varnothing} \cdot \frac{\exp(\kappa_M^e \cdot U_{d,M}^{e,e'})}{\exp(\kappa_M^e \cdot U_{d,M}^{e,\varnothing})}.$$

Similarly,

$$\frac{q_d^{e,e'}}{q_d^{\varnothing,e'}} = \frac{r_{d,F}^{e,e'}}{r_{d,F}^{\varnothing,e'}} = \frac{\exp(\kappa_F^{e'} \cdot U_{d,F}^{e,e'})}{\exp(\kappa_F^{e'} \cdot U_{d,F}^{\varnothing,e'})} \Rightarrow q_d^{e,e'} = q_d^{\varnothing,e'} \cdot \frac{\exp(\kappa_F^{e'} \cdot U_{d,F}^{e,e'})}{\exp(\kappa_F^{e'} \cdot U_{d,F}^{\varnothing,e'})}.$$

Combining the two gives us:

$$q_d^{e,e'} = [q_d^{e,\varnothing} \cdot \frac{\exp(\kappa_M^e U_{d,M}^{e,e'})}{\exp(\kappa_M^e U_{d,M}^{e,\varnothing})}]^{\frac{\frac{1}{\kappa_M^e}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}}} \cdot [q_d^{\varnothing,e'} \cdot \frac{\exp(\kappa_F^{e'} U_{d,F}^{e,e'})}{\exp(\kappa_F^{e'} U_{d,F}^{\varnothing,e'})}]^{\frac{\frac{1}{\kappa_F^{e'}}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}}}.$$

Taking logarithm on both sides, we have

$$\begin{aligned}
\log(q_d^{e,e'}) &= \frac{1}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}} \cdot (U_{d,M}^{e,e'} + U_{d,F}^{e,e'} - U_{d,M}^{e,\varnothing} - U_{d,F}^{\varnothing,e'}) + \frac{\frac{1}{\kappa_M^e}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}} \log(q_d^{e,\varnothing}) + \frac{\frac{1}{\kappa_F^{e'}}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}} \log(q_d^{\varnothing,e'}) \\
&\quad (\text{using equation (5)}) \\
&= \frac{1}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}} \cdot (\overline{V}_d^{e,e'} - \overline{V}_d^{e,\varnothing} - \overline{V}_d^{\varnothing,e'}) + \frac{\frac{1}{\kappa_M^e}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}} \log(q_d^{e,\varnothing}) + \frac{\frac{1}{\kappa_F^{e'}}}{\frac{1}{\kappa_M^e}+\frac{1}{\kappa_F^{e'}}} \log(q_d^{\varnothing,e'}).
\end{aligned}$$

C Quantitative Implementation

This section describes how we calibrate and solve the model.

C.1 Solving the Model

In solving the model, we take as given the following fundamental parameters: the exogenous components of city productivity and amenities $\{\bar{A}_d^e,\ \bar{K}_d^e\}$, local land supply shifter and elasticity, $\{\bar{H}_d,\ \epsilon_d\}$, the endowments of the four types of workers in each city $\{L_{o,s}^e\}$, the non-economic component of marriage surplus $\{\delta^{e,e'}\}$, the parameters governing idiosyncratic taste draws $\{\kappa_s^e, \theta_s^e, \eta^e\}$, gender biases and home production value $\{\bar{n}^e, \beta_F^e\}$, migration costs $\{d_{od}^e\}$, spillover elasticities $\{\gamma_{e,e'},\ \sigma_{e,e'}\}$, and utility function parameters $\alpha,\ \beta$.

Given these parameters, we solve for a set of prices and quantities, including migration decision π_{od}^e; the utility of different households $\{\overline{V}_d^{e,\varnothing}, \overline{V}_d^{\varnothing,e}, \overline{V}_d^{e,e'}\}$, the division of household utility between the spouses $\{U_{d,M}^{e,e'}, U_{d,F}^{e,e'}, U_{s,M}^{e,\varnothing}, U_{d,F}^{\varnothing,e'}\}$ and the expected utility of a city $\{\overline{U}_{d,s}^e\}$; the labor force participation decision of married women $\{l_d^{e,e'}\}$; total effective labor supply in a city $\{E_d\}$ and wage $\{W_{d,s}^e\}$; housing supply $\{H_d\}$ and rent $\{r_d\}$; government transfer t. We obtain these prices and allocation as solution to a system of equations, described below:

Problem C.1. *The following system of equations defines a solution to the competitive equilibrium of the model:*

1. *Equation (2): migration decision is optimal.*
2. *Equations (5), (8), (9) and (10): each local marriage market is a stable match.*
3. *Equations (13), (14), and (15), i.e., household utility and labor force participation decisions are consistent with local wages, rents, and amenities.*
4. *Equations (16) and (17): migration and labor supply are consistent with population and labor distribution*
5. *Equations (19), (22), and (23): amenities, productivity, and wags are consistent with individual decisions*
6. *Equations (24) and (25): the housing market clears*
7. *Equation (26): government budget balances*

C.2 Calibrating the Model

As summarized in panel C of Table 5, our calibration matches the level of rent by city, and the numbers and wages of high- and low-skill people by city, which pins down $\{\bar{H}_d\}$, $\{\bar{A}_d^e\}$, and $\{\bar{K}_d^e\}$, respectively. To calibrate the model, we find $\{\bar{H}_d\}$, $\{\bar{A}_d^e\}$, and $\{\bar{K}_d^e\}$, along with all endogenous variables described in Section C.1, as a solution to the following system of equations.

Problem C.2. *The following system of equations calibrates the competitive equilibrium of the model to the data:*

1. *All equations listed in Problem C.1*

2. *Average wage for skill type e in city d equals its data counterpart:* $W_d^e(\bar{K}_d^e) = \underbrace{\hat{W}_d^e}_{data}$

3. *Rent in city d equals its data counterparts:* $r_d(\bar{H}_d) = \underbrace{\hat{r}_d}_{data}$

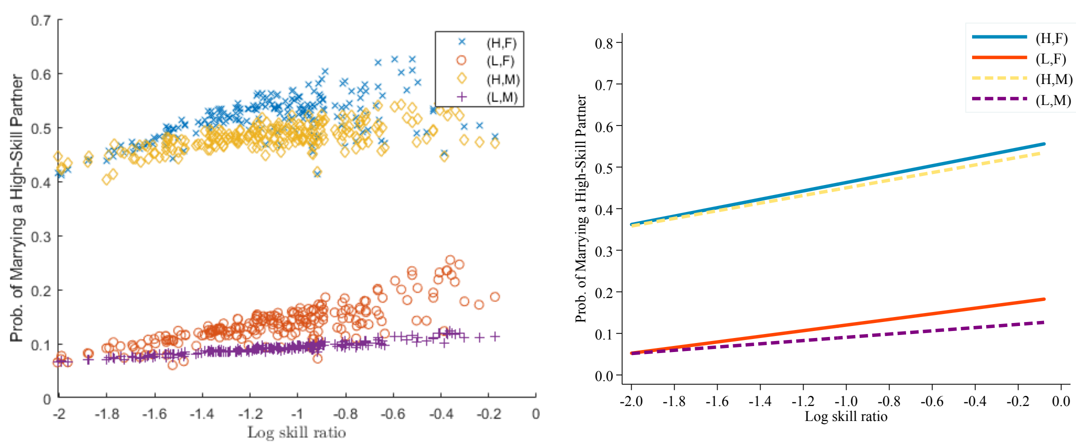
4. *The number of skill e people in d equals its data counterpart:* $\tilde{N}_d^e(\bar{A}_d^e) = \underbrace{\hat{\tilde{N}}_d^e}_{data}$.

In this problem, we write $W_d^e(\bar{K}_d^e)$, $r_d(\bar{H}_d)$, and $\tilde{N}_d^e(\bar{A}_d^e)$ as functions of fundamental parameters to highlight what identifies each set of parameters. The empirical counterparts of the model outcomes are denoted with a hat. Compared to Problem C.1, Problem C.2 has three additional sets of unknowns and three additional sets of equations.

To calibrate the model, we solve the system of equations listed in Problem C.2. Once we have obtained the solution, we can solve Problem C.1 for counterfactual equilibria.

C.3 Additional Results from Quantitative Experiments

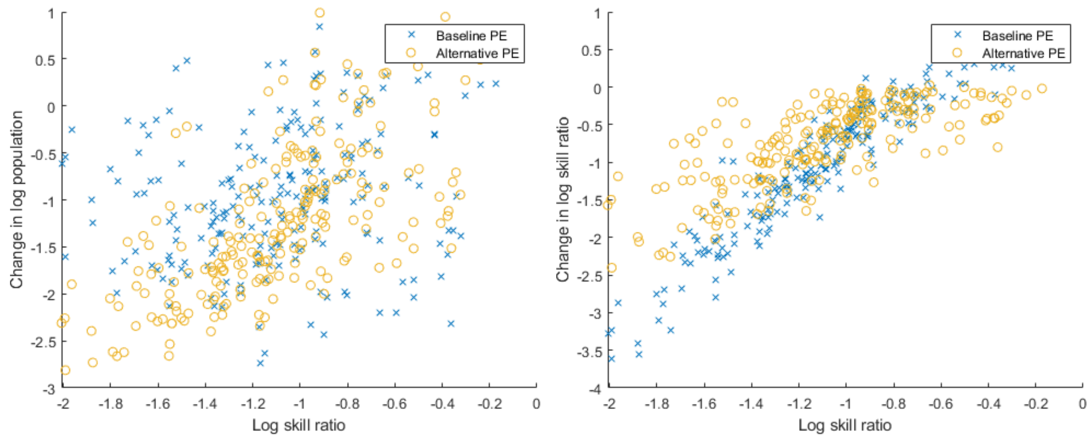
The model fits the probability of marrying a high-skill spouse. One might be concerned that our model might have failed to match that in big and skill intensive cities, people have a higher chance of marrying a high-skill spouse. If so, it could explain why marriages could be a force of dispersion despite being highly assortative in the data. Figure C.1 shows that this is not the case. The left panel of the figure is the model predicted probability of marrying a high-skill spouse as a function of the log skill ratio of cities; the right panel is the corresponding empirical estimates. The model aligns closely with the data.



(a) Model (b) Data

Figure C.1: Prob. of Marrying a Skilled Spouse: Data and Model

The left panel is the model's prediction on the probability of marrying a high-skill spouse. The right panel is the probability of an individual marrying a high-skill spouse in the data, see notes under Figure 1 for detail.



(a) City Size (b) Log Skill Ratio

Figure C.2: Comparison between Baseline and Alternative Partial Equilibrium Experiments
The figures compare the baseline partial equilibrium (PE) experiment from Figure 5 with an alternative PE experiment, in which we calculate the counterfactual expected utility under the assumption that the single rate of low-skill people in each city is the same as the single rate of high-skill people in the same city. Thus, in this alternative experiment, the different responses in high- and low-skill people is not due to their having different single rates.

Baseline versus alternative partial equilibrium experiments. We find that eliminating the marriage premia of cities in partial equilibrium increases the size and the skill share of currently skill-intensive cities. To understand what accounts for the differential responses of high- and low-skill people in this experiment, we conduct an alternative partial equilibrium experiment.

In this alternative experiment, we change the expected utility governing the migration decision of high-skill people in the same way as in the baseline experiment, i.e., by calculating their migration decision (equation (2)) assuming $\overline{U}^{H}_{d,s}$ is given by:

$$\overline{U}^{H}_{d,M} = \frac{\gamma}{\kappa^{H}_{M}} + \overline{V}^{H,\varnothing}_{d} - \frac{1}{\kappa^{H}_{M}}\log(1) = \frac{\gamma}{\kappa^{H}_{M}} + \overline{V}^{H,\varnothing}_{d}$$
$$\overline{U}^{H}_{d,F} = \frac{\gamma}{\kappa^{H}_{F}} + \overline{V}^{\varnothing,H}_{d} - \frac{1}{\kappa^{H}_{F}}\log(1) = \frac{\gamma}{\kappa^{H}_{F}} + \overline{V}^{\varnothing,H}_{d}.$$

We change the expected utility governing the migration decision of low-skill people to below:

$$\overline{U}^{L}_{d,M} = \frac{\gamma}{\kappa^{L}_{M}} + \overline{V}^{H,\varnothing}_{d} - \frac{1}{\kappa^{L}_{M}}\log(r^{L,\varnothing}_{d,M}) + \frac{1}{\kappa^{L}_{M}}\log(r^{H,\varnothing}_{d,M})$$
$$\overline{U}^{L}_{d,F} = \frac{\gamma}{\kappa^{L}_{F}} + \overline{V}^{\varnothing,L}_{d} - \frac{1}{\kappa^{L}_{F}}\log(r^{\varnothing,L}_{d,F}) + \frac{1}{\kappa^{L}_{F}}\log(r^{\varnothing,H}_{d,F}).$$

This experiment effectively calculates the hypothetical partial equilibrium migration decision for low skill people assuming their single rate in a city is the same as the high-skill people in the same city. Therefore it purges out the difference between the response of high- and low-skill people to the removal of marriage market premia due to their different single rates.

Figure C.2 compares the result from this alternative experiment to that from the baseline experiment. It shows that this alternative experiments generates a skill gradient that is only slightly weaker than in the baseline experiment. This suggests that the increase in skill concentration from the baseline equilibrium is not due to the differences in single rates across types, but rather due to different migration frictions and birth states.

C.4 Sensitivity Analysis

In this subsection we show that the main quantitative results are insensitive to external choice of β (the home good share) and the value of p_n (the price for the market alternative of home goods) in 1960.

We report two exercises. First, we consider different values of β in calibrating the model to the 2000 economy. We show that the choice of β does not matter for the quantitative impacts of eliminating marriages on spatial concentration from the 2000 economy. Second, we consider different values for p_n for 1960 (while always assuming $p_n = 1$ for 2000). We show that different choices do not affect the impact of the fundamental changes between 1960 and 2000 on spatial divergence.

Eliminating marriages on the basis of the 2000 economy. Table C.1 shows that given the observed female labor force participation rate, different choices of β lead to different inferred values for $\bar{n}^e$. However, eliminating marriages from these economies generate the same results, as summarized by the two gradient metrics in column 6 of Table 6.

Table C.1: Eliminating Marriages from the 2000 Economy

	Baseline ($\beta = 0.2$)	$\beta = 0.1$	$\beta = 0.3$
Calibrated $\bar{n}^e$			
$\bar{n}^H$	0.0043	0.0059	0.0034
$\bar{n}^L$	1.03	1.40	0.82
Results from counterfactuals			
The gradient of $\Delta \ln(pop)$ w.r.t. $\ln(\frac{H}{L})$	1.55	1.55	1.55
The gradient of $\Delta \ln(\frac{H}{L})$ w.r.t. $\ln(\frac{H}{L})$	1.75	1.75	1.75

Note: The table reports the main result under different choice of β. The upper panel reports the calibrated $\bar{n}^e$ corresponding to different values of β; the lower panel reports the counterfactuals from eliminating marriages.

The changes between 1960 and 2000. In Table C.2, we re-calibrate the model to match the 1960 data under different choices of p_n for 1960. Recalling that we normalize p_n to 1 in the 2000 economy, our choices in Table C.2 (p_n ranges from 1.5 to 3) thus implies a decrease in p_n between 1960 and 2000 that ranges between 33% to 66% (the baseline calibration assumes a 50% decease). The upper panel of the table shows that these choices lead to different values of $\bar{n}^e$ for 1960. The lower panel shows that, regardless of these values, once we feed the implied changes in p_n, $\bar{n}^e_h$ into the model, the model predicts the same increase in spatial divergence (see column 4 of Table 8).

Table C.2: Spatial Divergence between 1960 and 2000

	Baseline ($p_n = 2$ in 1960)	$p_n = 1.5$ in 1960	$p_n = 3$ in 1960
Calibrated 1960 $\bar{n}^e$			
$\bar{n}^H$	0.62	0.78	0.45
$\bar{n}^L$	2.19	2.76	1.58
Results from counterfactuals			
The gradient of skill	0.054	0.054	0.054
The gradient of population	0.20	0.20	0.20

Note: The table reports the calibrated $\bar{n}^e$ for 1960 under different choices of p_n for 1960. It also shows that despite the difference in inferred $\bar{n}^e$, alternative calibrations imply the same spatial divergence between 1960 and 2000.