2018

Aging And The Gains From Marriage

Toban Wiebe
University of Pennsylvania, tobanw@gmail.com

Aging And The Gains From Marriage

Abstract
Men and women have distinct marriage patterns over the lifecycle. In the contemporary USA, marriages for women are concentrated earlier in the lifecycle, whereas for men they are spread out later in the lifecycle. In particular, this means that men are more likely than women to get married in middle age and beyond. This difference is especially pronounced for remarriages — men are far more likely than women to remarry after the age of 30. As a result, there are far more single women than single men over the age of 40. This difference in remarriage patterns cannot be explained by the presence of children — in fact, among divorced women, those with children are more likely to remarry than those without. I investigate how the gains from marriage change over the lifecycle for men and women to understand whether these observed marriage patterns are driven by changes in the value of marriage, as opposed to being products of equilibrium sorting. I develop an equilibrium search and matching model that incorporates an aging process. This allows the model to capture both the lifecycle dynamics of marriage and divorce decisions as well as the impact of local population supplies on equilibrium matching outcomes. Using data from a large cross-sectional survey of the USA, I structurally estimate the model for 20 large city-level marriage markets. I recover an estimate of the gains from marriage, represented by a marital production function, in terms of the ages of husbands and wives. I find that marital output drops off twice as steeply with respect to female age, compared to male age. This suggests that women remarry less because the benefits are smaller, not just because of reduced availability of single men. Finally, I estimate a model in which people are characterized by their education and race in order to capture assortative mating along these dimensions. The results concerning age do not qualitatively change. I find large differences in marital output based on college attainment, but not race.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Petra Todd

Keywords
Aging, Marriage, Matching, Search

Subject Categories
Demography, Population, and Ecology | Economics

AGING AND THE GAINS FROM MARRIAGE

Toban Wiebe

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

____

Petra Todd, Edmund J. and Louise W. Kahn Term Professor of Economics

Graduate Group Chairperson

____

Jesús Fernández-Villaverde, Professor of Economics

Dissertation Committee:

Petra Todd, Edmund J. and Louise W. Kahn Term Professor of Economics

Kenneth Burdett, James Joo-Jin Kim Professor of Economics

Corinne Low, Assistant Professor of Business Economics and Public Policy

AGING AND THE GAINS FROM MARRIAGE

To my wife, Daweon.

Acknowledgments

I am very grateful to my advisor, Petra Todd, for her patience and support. Her guidance has been invaluable in the process of writing this dissertation. I am also grateful to my dissertation committee members, Kenneth Burdett and Corinne Low, for many productive discussions. I am especially indebted to Corinne Low for her excellent advice, which has helped me to develop as a researcher.

I would also like to thank several other faculty members. Andrew Shephard provided helpful comments on my research. I benefited from discussions with and suggestions from Jeremy Greenwood. I owe my inspiration for pursuing the topic of marriage matching to Andrew Postlewaite.

I also benefited greatly from discussions with Sophie Shin, Jan Tilly, and Edoardo Ciscato. Takeaki Sunada provided very helpful feedback on my research. Nitin Krishnan helped me in initially learning the workings of search and matching models.

I am grateful to my classmates, for their willingness to help one another and for their friendship. I thank my family for their understanding and support. Finally, I thank my wife, Daweon Ryu, for her unfailing encouragement.

Toban Wiebe

Philadelphia, PA

July 1, 2018

ABSTRACT

AGING AND THE GAINS FROM MARRIAGE

Toban Wiebe

Petra Todd

Men and women have distinct marriage patterns over the lifecycle. In the contemporary USA, marriages for women are concentrated earlier in the lifecycle, whereas for men they are spread out later in the lifecycle. In particular, this means that men are more likely than women to get married in middle age and beyond. This difference is especially pronounced for remarriages — men are far more likely than women to remarry after the age of 30. As a result, there are far more single women than single men over the age of 40. This difference in remarriage patterns cannot be explained by the presence of children — in fact, among divorced women, those with children are more likely to remarry than those without. I investigate how the gains from marriage change over the lifecycle for men and women to understand whether these observed marriage patterns are driven by changes in the value of marriage, as opposed to being products of equilibrium sorting. I develop an equilibrium search and matching model that incorporates an aging process. This allows the model to capture both the lifecycle dynamics of marriage and divorce decisions as well as the impact of local population supplies on equilibrium matching outcomes. Using data from a large cross-sectional survey of the USA, I structurally estimate the model for 20 large city-level marriage markets. I recover an estimate of the gains from marriage, represented by a marital production function, in terms of the ages of husbands and wives. I find that marital output drops off twice as steeply with respect to female age, compared to male age. This suggests that women remarry less because the benefits are smaller, not just because of reduced availability of single men. Finally, I estimate a model in which people are characterized by their education and race in order to capture assortative mating along these dimensions. The results concerning age do not qualitatively change. I find large differences in marital output based on college attainment, but not race.

Contents

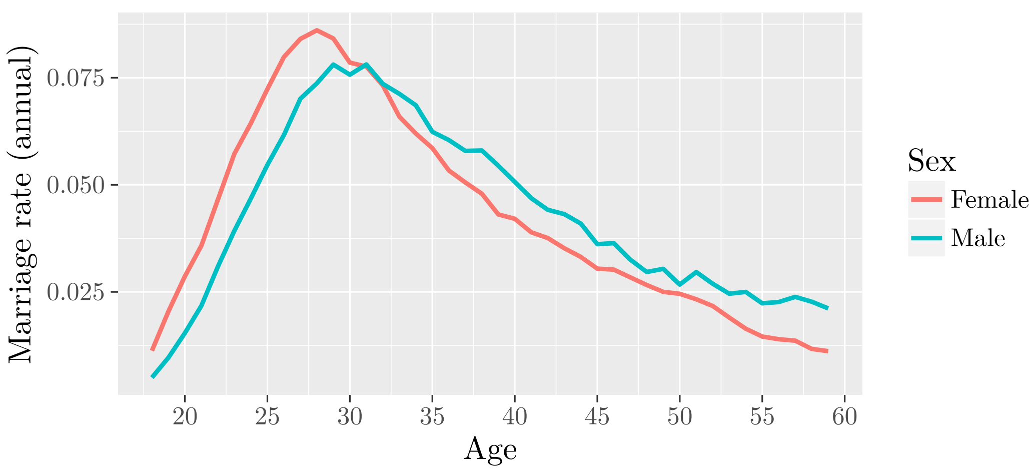
List of Tables

List of Figures

Chapter 1

Introduction

Male and female marriage rates follow distinct patterns over the lifecycle. In the contemporary USA, female marriage rates are higher below the age of 30, above which male marriage rates dominate. The higher male marriage rate at later ages is in large part due to a much higher male remarriage rate (as remarriages account for larger shares of all marriages at later ages).



Notes: Marriage rates are computed as the proportion of all singles who enter marriage in the survey year. ACS 2010–2014.

Figure 1.1: Marriage rate gap by age

Figure 1.1 shows the annual marriage rates of women and men at each age. Female marriage rates rise earlier and peak at age 28. Male marriage rates rise later and peak around age 30 — the peak is lower but drops off more slowly, so that by their mid-30s, men have substantially higher marriage rates. This pattern of earlier female marriage is broadly explained by differential fecundity:[1] the female biological clock ticks faster, so women have stronger incentives to marry earlier. Insofar as part of the value of marriage comes from raising children, a woman's value as a partner is tied to her fecundity. Low (2015) calls this "reproductive capital", as the potential to bear children can be thought of as a depreciating economic asset. She uses an online dating experiment to estimate the "price" of fecundity, finding that men value an extra year of fecundity on par with an additional $7,000 of annual income from a prospective partner.

As men are effectively not bound by this reproductive constraint,[2] they can afford to wait longer before marriage, taking extra time to search for a better match and also to accumulate resources that will improve their standing in the marriage market. Whether differential fecundity alone explains this difference or not,[3] it is clear that aging affects women and men differently on the marriage market. I refer to this phenomenon more generally as "differential aging".[4]

In this paper, I abstract away from the specific mechanisms underlying differential aging and instead focus on quantifying the value of marriage in terms of the ages of the husband and wife. This provides insights into how age, as an individual characteristic, affects marriage and divorce patterns for men and women over the lifecycle. In particular, I estimate the gains from marriage as a function of both spouses' ages in order to understand the sex

[1] The term fecundity refers to reproductive potential whereas fertility refers to realizations of that potential, although in the common usage fertility also refers to reproductive potential. Siow (1998) shows how differential fecundity can account for many of the observed sex differences in marriage and labor, including marriage timing and the spousal age gap.

[2] Male fecundity also declines with age, but much more slowly. Men are able to father children well beyond the age of female menopause.

[3] A complementary interpretation of the differential fecundity mechanism is that men care about fecundity indirectly via its correlation with youth and hence beauty. This is because mating preferences (i.e., the sense of beauty) would have been shaped by strong evolutionary pressures to favor fecund mates.

[4] Differential aging could also arise for reasons other than fecundity. For example, it may be the case that men mature later in terms of their suitability for marriage and fatherhood. This could be due to their psychological disposition and/or to their ability to earn to support a family. These factors would affect a man's marriageability in the same sense that fecundity affects a woman's.

differences in marriage patterns seen in Figure 1.1.

The distribution of marriages and of singles provides evidence about the gains from marriage for each age or type pair. One may reason that, by revealed preference, more common pairings produce greater marriage gains and vice versa. For example, most married couples are of similar ages, which suggests that the value of marriage is greatly reduced when there is a large age difference. However, this mode of inference will be confounded by equilibrium sorting in the marriage market. As a one-to-one matching setting, the equilibrium of a marriage market follows a supply-and-demand logic.[5] A sex ratio imbalance will improve the matching outcomes for the sex in short supply, whereas the opposite sex will face greater competition and some will be excluded from matching at all. Thus, equilibrium sorting will affect the matching outcome in ways that do not reflect the underlying gains from marriage. Observed matching outcomes reflect both preferences and equilibrium sorting. For instance, the fact that husbands are typically slightly older than their wives does not necessarily imply that such marriages are more productive than same-age marriages or marriages in which wives are older than husbands. Instead, it could just be the result of a relative scarcity of marriageable younger men, resulting in a matching equilibrium in which women marry up in age.

Although males and females are born in roughly equal numbers, the relative population supplies of men and women in local marriage markets can vary due to migration, mortality, and incarceration. Taking social groups into consideration, there are even more ways that the relevant sex ratios might vary. Table 1.1 shows the extent to which marriages are homogamous with respect to education and race.[6] For example, among those who are married, 66.4% of college-educated women and 77.4% of college-educated men have a college-educated spouse — under random matching, these numbers would be 36.4% and 42.5%,

[5]This is easily seen in matching models such as Gale and Shapley (1962), Becker (1973), or Shimer and Smith (2000). A number of papers have empirically studied the impact of sex ratio imbalances on marriage outcomes. For example, Angrist (2002) and Abramitzky, Delavande, and Vasconcelos (2011) both find large effects on marriage probabilities, consistent with the predictions of matching models.

[6]Greenwood et al. (2014) documents rising levels of positive assortative mating with respect to education since the 1970s. Regarding low interracial marriage rates, Wong (2003b) investigates whether social norms have been a deterrent, and Shin (2014) examines the extent to which limited meeting opportunities can account for the shortfall.

Group	Women	Men
Non-college	85.2% (63.6%)	77.8% (57.5%)
College	66.4% (36.4%)	77.4% (42.5%)
White	97.1% (91.8%)	97.1% (92.1%)
Black	88.5% (8.2%)	77.1% (7.9%)

Notes: Values are computed as the percentage of women/men in homogamous marriages out of all married women/men in the given group. The hypothetical values under random matching are shown in parentheses. The basis is calculated separately for women and men. To calculate the values for women, the basis is all marriages in which the wife is at most 40 years old, and for men, the basis is all marriages in which the husband is at most 42 years old. ACS 2008–2016.

Table 1.1: Homogamous marriage rates versus random matching baseline

respectively. If these high levels of assortative matching reflect marital preferences, then the sex ratios within these groupings become important factors in the equilibrium outcome. It is well known that women complete college at higher rates than men.[7] This means that college-educated women face a shortage of similarly-educated men. To the extent that they prefer to marry college-educated men, this worsens their opportunities in the marriage market.

1.1 Marriage market models

In order to account for equilibrium sorting, I estimate an equilibrium matching model in order to infer the gains from marriage over the lifecycle. There are two main classes of empirical marriage market frameworks that are used to estimate the gains from marriage: (1) the static matching framework of Choo and Siow (2006), and (2) the search and matching framework of Goussé, Jacquemet, and Robin (2017b) which is based upon the model of Shimer and Smith (2000).[8] In the static matching framework, the solution concept is that of

[7]Women have completed college in greater number than men since the 1980s. Bronson (2015) reports that women now make up 58% of college graduates. She argues that, as divorce rates jumped in the 1970s, women sought college degrees as a form of insurance against very low income in the event of divorce.

[8]Shimer and Smith (2000) is the canonical search and matching model with transferable utility. It is a model of one-to-one matching and can be applied to two-sided matching markets, such as marriage and job search.

stable matchings, and there is no notion of time. The matching model elegantly incorporates the multinomial logit choice structure,[9] which yields a simple closed-form expression to non-parametrically estimate the gains from marriage.[10] In the search framework, agents face search frictions and so it takes time to find a partner. The solution concept is that of stationary (i.e., steady-state) equilibrium.[11] Both frameworks assume large markets (continuous populations) to avoid problems of discreteness.

These two frameworks share a fundamental similarity in that they both assume transferable utility within marriage in order to non-parametrically identify the underlying gains from marriage — the extra utility generated within marriage relative to being single. As Choo and Siow (2006) explain in their Introduction, transferable utility is the key assumption that allows for the non-parametric identification of the gains from marriage. This is because transferable utility obviates the need to estimate separate marital preferences for each sex, as it suffices to estimate the joint marital output.[12] For example, in a model with I types of men and J types of women, there are $I \times J$ possible type pairs, and so each sex has potentially $I \times J$ match preference parameters. If the researcher only observes the number of singles and pairs of each type, this provides $I + J + I \times J$ observations, which for $I, J > 2$ is strictly less than $2 \times I \times J$, the number of potential preferences parameters. Thus, identifying assumptions are necessary for any empirical matching model. With the assumption of transferable utility, there are at most $I \times J$ parameters describing the joint marital output, and so it can be identified. With non-transferable utility, other identify-

[9] Galichon and Salanié (2015) generalize the model beyond the multinomial logit structure, which imposes restrictive substitution patterns. The generalized model can accomodate more realistic distributions of unobserved heterogeneity, e.g., correlations in unobserved preferences over different types.

[10] In the static matching framework, men and women are characterized by types. Preferences over potential matches (i.e., opposite-sex types) are represented by two additive components: a systematic component common to everyone of that type, and an idiosyncratic component that is modeled as a random preference shock.

[11] In the search and matching framework, men and women may also be characterized by types. Furthermore, married couples may separate and return to the pool of singles. Preferences over potential matches are derived from a marital production function, which is defined over types, and an additive idiosyncratic match-specific bliss shock that is realized upon meeting. Thus, as with the static matching case, match preferences are determined by a systematic component and an idiosyncratic component.

[12] Under transferable utility, couples behave as a single decision-maker, and so each couple aims to maximize the joint marital output. This also means that stable matchings maximize the sum of marital output in the society.

ing assumptions would be required in order to estimate separate preferences for each sex. Typically, parametric restrictions are imposed on the utility functions.

In both frameworks, the gains from marriage are represented by a marital production function whose output is divided between the two spouses.[13] As a model primitive, it captures the fundamental gains from marriage, independent of the particular marriage market equilibrium. Another convenience provided by the transferable utility assumption is that the reduced form of the models does not include equilibrium transfers (i.e., prices). As a result, both frameworks allow for estimation without requiring a full solution of the model, and so they are relatively simple to implement.[14] By assuming that the data come from a matching equilibrium (a stable matching or a stationary equilibrium), these models are identified from a single cross-section of the population. The static model requires only the population stocks of marriages and singles, whereas the search model additionally requires the corresponding flows into and out of marriage.[15]

To study lifecycle marriage patterns, it is crucial that the model be dynamic and incorporate an aging process so that agents optimize over the lifecycle. The incorporation of lifecycle dynamics into the static matching framework of Choo and Siow (2006) has been made by Choo (2015), who embeds the static matching model into a discrete choice dynamic programming framework. In this overlapping-generations model, a frictionless marriage market clears in each period, with those who opted to remain single going on to participate in the marriage market in the next period. Men and women are characterized by their age,

[13]In the static matching framework, the output is divided according to the supply and demand for spouses on the marriage market, with transfers being analogous to market prices. In particular, the equilibrium prices clear the market and there is no surplus to bargain over. In the search framework, by contrast, search frictions give rise to bilateral rents, as it takes time to find potential matches. As such there is scope for bargaining over the resulting surplus. The convention in the literature, e.g., Shimer and Smith (2000), is to divide the surplus according to Nash bargaining, with the outside options of being single and searching as threatpoints.

[14]Non-transferable utility models can be used to parametrically estimate marital preferences, but they require that the model be solved in order to be estimated. In practice, solving for an equilibrium must be done numerically. Wong (2003a) estimates a search model with non-transferable utility, based on the model of Burdett and Coles (1997). Coles and Francesconi (2017) generalizes the search framework to a collective household model that nests the special cases of transferable and non-transferable utility.

[15]Flows can be obtained from cross-sectional datasets with information on whether the respondents entered or exited marriages during the past year. Search models are often estimated from panel data using a durations-based maximum likelihood estimator (e.g., Goussé (2014), Shin (2014), and Wong (2003a)), but this severely restricts sample sizes as compared to using cross-sectional datasets.

which increases in each period. To replenish the population as it ages, a new generation of young people enters the marriage market in each period. The model assumes a stationary equilibrium in order to identify the gains from marriage by age. As with the model of Choo and Siow (2006), it can be estimated without needing to solve the model, and can still be identified from just the population stocks of singles and couples by age. However, rather than adopting this model, I opt instead for a search model, which additionally incorporates information on population flows into and out of marriage.

For the purpose of estimating the gains from marriage over the lifecycle, the search framework provides several distinct advantages. Most importantly, it can parsimoniously accomodate an endogenous divorce process.[16] By contrast, Choo (2015) models divorce as an exogenous shock, acknowledging that "[a] cost of this approach is that the model has nothing to say about the division of within-marriage surplus over the life-cycle". Related to this, the estimated gains from marriage have a slightly different interpretation: in my search model, the marital production function represents a flow of output (gains from marriage) at specific ages, and so it changes over the lifecycle as the couple ages; for Choo, the marriage gains represent the discounted present value of the utility from entering into a marriage today. Estimating the age-specific flow utility enables me to see how the gains from marriage change over the course of a marriage. This is of course very important for understanding divorce and remarriage choices. Finally, Choo (2015) studies how marriage gains changed between 1970 and 1990 and how a static model underestimates the gains for the young, whereas I use recent data to study sex differences in the gains from marriage and how they change over the lifecycle.

Beyond Choo (2015), a few other papers have incorporated an aging process into empirical equilibrium marriage market models, though these do not provide for the non-parametric estimation of the marriage gains as with the two frameworks discussed above. Instead, they estimate preference parameters by means of the Simulated Method of Moments (SMM). In

[16] I follow the approach of Goussé (2014), who developed the endogenous divorce process by introducing periodically updating match quality shocks, which may trigger a decision to separate. With transferable utility, there is no need to make any assumption about commitment in marriage, as couples choose to marry or divorce if and only if the decision is mutually beneficial.

a related paper, Rıos-Rull, Seitz, and Tanaka (2016) employ a simple stochastic aging process in a discrete-time search model of marriage in order to study the effects of sex ratios on marriage patterns. They fit the model to 1950 data and show that sex ratio effects on equilibrium sorting can account for much of the changes in marriage patterns since 1850. They estimate preference parameters over spousal age (utility is non-transferable) as well as parameters governing the rate at which each sex matures, capturing the notion of differential aging. Their results are supportive of the differential aging hypothesis: the parameter estimates suggest that women mature earlier than men, but that men remain attractive in the marriage market for a longer portion of their lives.

Díaz-Giménez and Giolito (2013) explicitly model differential fecundity in a discrete-time search model by basing the gains from marriage partly on having children and calibrating the lifecycle fecundity profiles of men and women according to estimates from the medical literature. They find that the age gap in marriage can be explained by differential fecundity, but not by sex differences in income as has traditionally been proposed. Bronson and Mazzocco (2018) incorporate a simple differential aging mechanism into their discrete-time search model by imposing the restriction that women can only marry when young, whereas men can marry when young or old. They use the model to show how changes in marriage patterns over time can be explained by a combination of differential aging and sex ratio effects. Finally, Coles and Francesconi (2011) incorporate continuous aging into a search model and provides theoretical results showing that "toyboy" marriages — between younger, poorer men and older, richer women — can arise as an exchange of youth for wealth. Unfortunately, using a continuous aging process makes it much harder to solve the model (see Appendix E for a discussion of the difficulties that arise).

1.2 Estimating the gains from marriage over the lifecycle

I use the framework of Goussé, Jacquemet, and Robin (2017b) to develop a search model that incorporates an aging process in order to estimate the gains from marriage in terms of the ages of each spouse.[17] I derive results for solving, estimating, and numerically simulating the model. The model also features multidimensional types (in addition to age) and differential mortality by sex, age, and type.

As marriage markets are circumscribed more closely to the city level than to the nation or even state level, I define marriage markets at the level of the Metropolitan Statistical Area (MSA). The definition of an MSA is intended to capture the notion of an economic region around a city, as measured by commuting and employment. This is also a fitting geographical delineation of a marriage market. To my knowledge, no other paper estimates a marriage market model at a plausible geographic scale — typically, sample size constraints require that marriage markets be defined at the national level.[18]

As each marriage market has its own equilibrium, aggregating markets to a larger geographic area such as a country comes at the price of losing the heterogeneity of local markets.[19] This is a serious practical concern, as sorting of men and women between cities can result in large differences in local sex ratios across marriage markets, even if the global sex ratio is balanced. For instance, the clustering of industries into particular cities can result in local labor markets with imbalanced sex ratios. New York City is a major fashion industry hub that employs many women, whereas Silicon Valley is a major technology industry hub that employs many men. It is obvious even to the casual observer that these two places have very different marriage markets. Moreover, rural industries such as farming and petroleum primarily employ men.

I treat local marriage markets as isolated from one another, each with its own equilibrium. However, I assume common model primitives which I estimate by using data from

[17] The framework is presented in its essential form in Goussé (2014), who applies the model to study marriage patterns in terms of wealth and beauty by estimating their contributions to the marital production function for each sex. I provide an outline of the model in Appendix D.

[18] An exception is Gayle and Shephard (2018), who recognize the value of defining marriage markets more locally. They estimate a marriage matching model at the Census Bureau division level (the US is partitioned into 9 such divisions). They note: "We do not use a finer level of market disaggregation due to sample size and computational considerations."

[19] I ignore the endogeneity of migration decisions and instead treat migration flows as exogenous. However, this is potentially an important factor for large cities, as their labor markets attract many young people to migrate. Such migration decisions depend on both labor and marriage markets, among other factors. Allowing for endogenous migration decisions would unify the different local marriage markets under one global equilibrium.

several local marriage markets.To attain the requisite sample size to make this possible, I pool several years of recent data from the *American Community Survey* (ACS) to estimate the model for several large MSAs.

By examining the estimated gains from marriage, I find that the marital output for a typical couple over 30 falls twice as rapidly in the wife's age as in the husband's age. Specifically, the partial derivative of the marital production function with respect to the wife's age is twice that for the husband's age. This indicates that differential aging plays a fundamental role, generating asymmetric gains to marriage in terms of aging for each sex. However, it does not appear that differential fecundity in particular is a major contributor to differential aging, although this may be due to the fact that the model does not account for children. I also find very large differences in marital output with respect to education. College graduates experience much higher gains from marriage, rationalizing their high marriage rates. This finding is consistent with other work on the marriage outcomes of different educational classes. By comparison, differences between races are relatively small, although minorities without college degrees experience substantially lower gains from marriage at later ages. This suggests that the lower marriage rates of minorities are primarily due to their large sex ratio imbalances.

1.3 Contribution to literature

This paper makes several contributions to the literatures that study marriage choice. First, I document a number of facts about male and female marriage patterns over the lifecycle. I show that women are more likely than men to enter marriage before age 40, whereas men are much more likely than women to remarry at later ages. This results in a highly unbalanced sex ratio for older people, with single women far outnumbering single men.

Second, I make a technical contribution to the marriage search framework. In their conclusion, Goussé, Jacquemet, and Robin (2017b) outline several directions for future research, noting that "our description of matching can and should be improved by introducing aging and the life cycle in the analysis." The present paper makes a key step in this direction, by extending the marriage search framework to allow for agent types to evolve through aging.[20] With an aging process, my model can capture the lifecycle dynamics of marriage timing decisions. For example, marriage rates for younger age groups may appear low not because the gains from marriage are smaller at those ages, but because the value of search is larger due to having a longer horizon over which to enjoy the gains from marriage. Additionally, the model accounts for the equilibrium effects arising from coordination on marriage timing — search is valuable when there are plentiful opportunities to meet other singles, but not so much once most people in the relevant age group have married off. As people within a given age cohort marry off and the pool of singles shrinks, the value of being single and searching falls. Anticipating this, singles become more willing to marry in the present, to avoid being "left on the shelf".

Third, this paper contributes to the literature that investigates sex differences in marriage timing, divorce and remarriage, and the consequences of differential fecundity. I estimate the gains from marriage for each age pair, which provides insight into how changing gains from marriage over the lifecycle affect the marriage market equilibrium. I find that female aging reduces marital output about twice as much as male aging does, providing empirical support for the differential aging hypothesis. However, I do not find any evidence that this difference is driven by differential fecundity, although this effect may be masked due to a limitation of the model. I also incorporate education and race as agent types into the matching model, and find that college graduates experience exceptionally large gains from marriage, whereas race plays only a minor role. This rationalizes the much higher marriage rates of college graduates, and suggests that sex ratio imbalances are an important factor in explaining the lower marriage rates of minorities. Finally, by performing my estimation at the local marriage market level, I account for the fact that different cities have different marriage market equilibria. By doing so, I obtain more credible model estimates as compared to the existing literature which typically assumes a single national marriage

[20] Existing applications of the Goussé, Jacquemet, and Robin (2017b) framework treat agent types as static and ignore aging. In these models, agents are ageless but may randomly die, e.g., Shin (2014) or Goussé, Jacquemet, and Robin (2017a). Other marriage search models such as Wong (2003a) and Coles and Francesconi (2017) do likewise.

market.

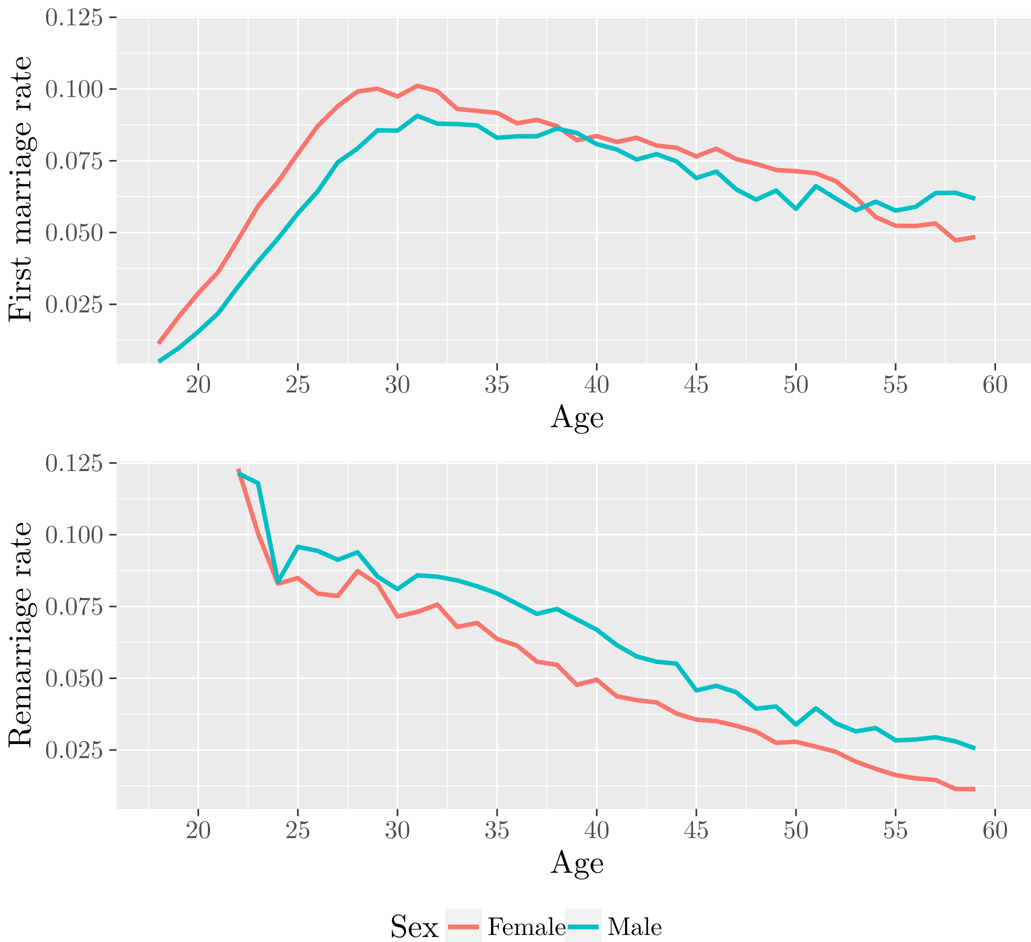
The rest of the paper proceeds as follows. Chapter 2 explores the lifecycle marriage patterns of men and women and discusses the literature on the topic. Chapter 3 introduces the lifecycle search model of marriage. Chapter 4 develops the identification strategy and derives the non-parametric estimation of the model primitives. Chapter 5 discusses the estimation results. Chapter 6 concludes.

Chapter 2

Data and Facts

Figure 2.1 displays marriage rates by age for first marriages and remarriages. The top panel shows that women have much higher first marriage rates up until their mid-30s, after which the gap closes. The bottom panel shows that men have higher remarriage rates, with a substantial gap that remains stable after the age of 30. Taken together, this means that women marry at younger ages (necessarily to older men than themselves on average), whereas men marry later and have much greater success in remarriage.

Another point to note is that, although male remarriage rates are decisively higher, male first marriage rates remain lower until the age of 54. The gap does narrow substantially starting around age 30, and rates even equalize in the latter half of the 30s. All else equal, a higher male remarriage rate necessitates a lower male first marriage rate as a matter of accounting: there is a fixed supply of single women available to marry, and so male remarriages must displace some male first marriages. The fact that men remarry more than women indicates that there is some serial polygyny in the marriage market, with some men marrying and remarrying multiple times and some men marrying late or not at all. This reflects a general principle from evolutionary biology: in mammals, the womb is the limiting factor in producing offspring and so fecund females are a scarce resource. As such, males compete to mate with (relatively scarce) fecund females, and so the fittest

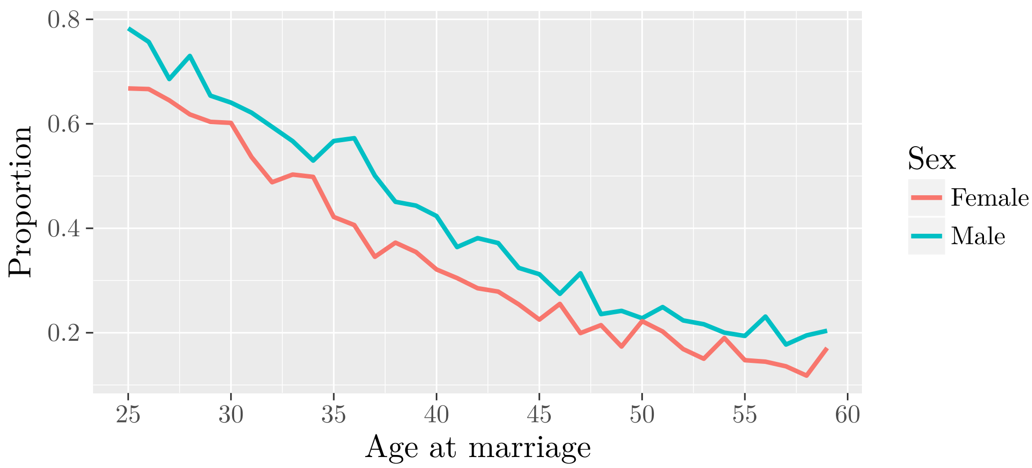


Notes: The first marriage rate is the proportion of all never-married singles who enter into marriage during the survey year. The remarriage rate is the proportion of all previously married singles who enter into a marriage during the survey year. ACS 2010–2014.

Figure 2.1: First marriage and remarriage rates by sex

males will get to reproduce more while some other males will not reproduce at all.[21] In this case, the least marriageable men will be crowded out of the marriage market by more desirable divorced men who marry multiple times.

Aside from differential mortality or migration, the only way that there can be a gap in remarriage rates is if divorced men disproportionately get remarried to previously never-married women. Figure 2.2 shows that, at every age, remarriages for men are more likely to involve a wife entering her first marriage than vice versa.



Notes: The age axis denotes the age of the person getting remarried. ACS 2010–2014.

Figure 2.2: Proportions of remarriages to previously never-married spouse

Figure 2.3 shows the distribution of the population and of singles in the four largest marriage markets. Each of these MSAs sees a similarly sharp decline in the number of singles from ages 25 to 35, which levels out at around age 40. The population sex ratio is most skewed in New York City, with women outnumbering men at every age above 24 and the gap widening further at later ages. For singles, each MSA follows the same overall sex

[21] Trivers (1972) explains how differential parental investment in offspring governs the mating strategies of males and females. Female animals typically invest far more in the production of offspring, such as by producing eggs or gestating offspring in the womb. As such, females face a higher opportunity cost in the choice of whom to mate with. As a result, they are much more selective than males, whose opportunity cost is practically negligible.

For species in which males do not make any parental investment, e.g., the lion or the hippopotamus, extreme polygyny is the rule — females will all seek to mate with the best available male. The females often live in a group along with this "alpha" male, and all other adult males are excluded. The alpha male is replaced when another male challenges and defeats him.

ratio pattern: single men outnumber single women at every age below 35, but past age 40, the number of single men drops below the number of single women. At later ages, the sex ratio among singes becomes severely imbalanced, approaching a 2:1 ratio. Not only does the difference in counts widen dramatically, but the sex ratio is further amplified because the supply of singles (the denominator) is much smaller by that age.[22] This reflects the two patterns observed in Figure 2.1: women marry earlier than men, but older men remarry more and are thus less likely to be single.[23]

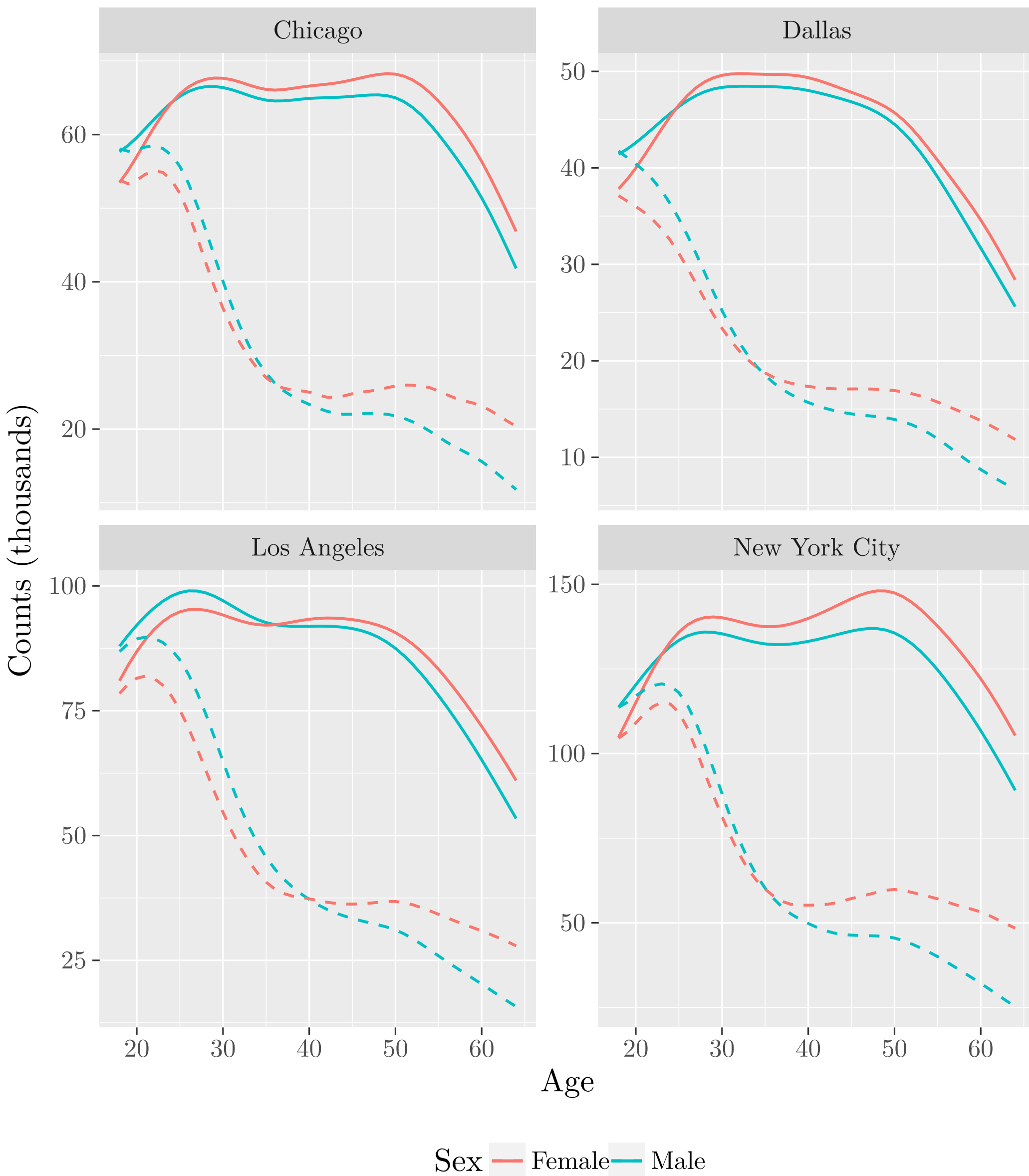
Taken together, these facts strongly suggest that, as they grow older, men enjoy greater success in the marriage market than do women. Male remarriage also contributes to the relatively higher female first marriage rates. If divorced men are getting remarried to younger, never-married women, this increases the first marriage rate for these younger women and decreases the remarriage rate for older women as well as the first marriage rate for the young men who are "crowded out". Another consequence of favorable remarriage opportunities for men is that married men have better outside options, which improves their bargaining position within marriage, and plausibly induces more divorce on the margin. As such, as women age, they not only face a tighter remarriage market, but their share of marital output shrinks as well.

In addition to the concerns about poverty among single mothers, there is another reason to worry about the worse marriage market opportunities of older women. Women face a more exacting tradeoff between family and career, and typically reduce their work hours for several years in order to allocate more time to raising their children. The result is that mothers gain less experience and human capital, resulting in lower earnings relative to their husbands. In the event of divorce, mothers often take custody of their children, further adding to their financial burden.

Bronson (2015) argues that women anticipate this and factor it into their career deci-

[22]Note that the overall population drops sharply past age 50. This drop is driven mainly by married couples leaving the city — the married populations shrink almost at almost double the rates of the single populations.

[23]These sex ratio patterns also hold when the population is segmented by education and race. Figures F.2 to F.5 in Appendix F show the population distributions by college attainment and race. Though there are differences between education and race groups, the qualitative pattern still holds, with many more women being single at later ages.



Notes: Counts are smoothed as described in Section 4.2. The solid lines denote total population counts whereas the dotted lines denote counts of singles. The y-axis is cropped from below for clarity. ACS 2008–2016.

Figure 2.3: Population distribution in largest MSAs

sions in two ways. First, women are more likely to get a college degree, she argues, as a form of insurance against very low earnings in case of divorce. Second, these women disproportionately select into more flexible but lower-earning majors such as education or nursing, as opposed to engineering or business. This way, they can more easily manage both motherhood and career. And by not leaving the workforce to raise children, these mothers can rely on their careers to provide for themselves in the event of a divorce. Poor remarriage prospects only amplify this mechanism, as the option to remarry for financial support becomes less viable. By this line of reasoning, weaker remarriage prospects for women contribute to the gender gaps in higher education and occupational choice.

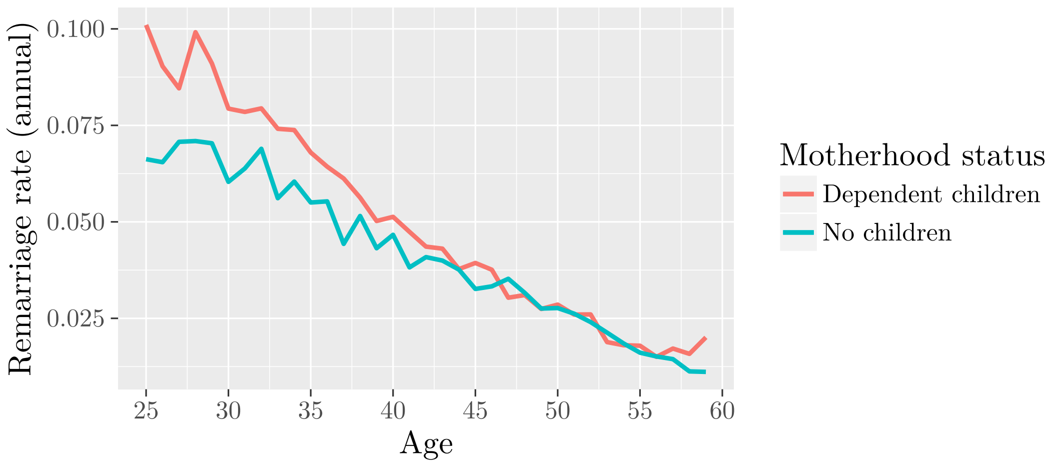
Below, I explore several plausible explanations for why men gain this advantage on the marriage market from their 30s onward. These explanations are not mutually exclusive and there is likely to be some interplay between them.

2.1 Motherhood and children

The fact that mothers typically take custody of their children after divorce could potentially explain why women are less likely to remarry. First, it could simply be that divorced mothers are busy with both work and taking care of their children, and do not have as much time for dating. Second, the prospect of step-parenthood may reduce their desirability as prospective spouses, as men may prefer to raise their own biological offspring. For obvious evolutionary reasons, animals are well-adapted to avoid investing resources in genetically unrelated offspring.[24] By this line of reasoning, the presence of step-children may substantially reduce the marital surplus as compared to couples raising their own biological offspring.

However, it is simply not the case that motherhood reduces marriage rates — Figure 2.4 shows that, among divorced women, mothers are actually more likely to remarry up until their mid-forties, and are no less likely to remarry thereafter. This could be because single

[24] A standard reference is Trivers (1972), which outlines the evolutionary pressures that determine parenting behavior in sexually-reproducing species.



Notes: Mothers are defined by the presence of children under 19 years of age in the home. ACS 2010–2014.

Figure 2.4: Female remarriage rates by motherhood status

mothers are more willing to get married or remarried. It could also be due to selection into motherhood — the more desirable women may be disproportionately represented among mothers. In any case, motherhood cannot be generating the remarriage gap.

2.2 Differential aging

Several papers have sought to explain the marital age gap — the fact that husbands are on average older than their wives — as a product of the shorter female reproductive horizon. Two such papers are Siow (1998) and Díaz-Giménez and Giolito (2013). By taking reproduction as the impetus for marriage, these models show how differential fecundity results in an equilibrium in which husbands are older than their wives. The underlying mechanism is that female desirability (in this case, fecundity) depreciates more rapidly with age than does that of males. Bronson and Mazzocco (2018) incorporate this mechanism into their marriage market model by allowing marriage only for young women but imposing no such restriction on men.

Low (2015) provides experimental evidence demonstrating that men care about fecundity directly (independently of age) in the dating market, and shows how women face a tradeoff between marriage and career because of their depreciating reproductive capital. Related to this idea, Dessy and Djebbari (2010) show how a coordination failure between younger and older women on the marriage market penalizes women who postpone marriage to focus on their careers.

Differential aging may also occur at the transition to adulthood — the sexes may mature at different rates in terms of their readiness for marriage. If men mature later in terms of their suitability for marriage and fatherhood, then young women who are ready for marriage will look to date older, more mature men. This could be due to mens' psychological disposition and/or to their ability to earn to support a family. Rıos-Rull, Seitz, and Tanaka (2016) allow for aging and age preferences in their marriage market model, and finds that women mature earlier than men in terms of their desirability in the marriage market. Their desirability also starts to decline at an earlier age than that of men.

To illustrate the differential aging mechanism, consider a simple matching model in which there are two traits, beauty and wealth. Suppose that men care more about beauty and women care more about wealth in their prospective spouses. Further suppose that wealth is non-decreasing in age, while beauty is strictly decreasing in age. Under these conditions, though aging reduces both mens' and womens' value on the marriage market, women face a greater age penalty because of mens' greater attention to beauty. In contrast, mens' value fares less badly (or even better) with age because of the wealth bonus as well as a smaller age penalty. As such, older men are better able to compete against young men for marriages to young women. Thus, older men will have higher marriage rates than women of the same age. Along these lines, Coles and Francesconi (2011) develop an equilibrium search model in which mate value depreciates over time, but there is also a countervailing career incentive to delay marriage. This leads to an equilibrium with matching between older high earners and younger low earners.

This differential aging mechanism can explain how sex differences in marriage rates will change with age. It predicts that, because of the greater female age penalty, older women will face more intense competition from younger women (whereas the reverse occurs for men), and hence female marriage rates will start to decline at earlier ages than male marriage rates. Thus, female marriage rates will peak and decline earlier than male marriage rates. Male marriage rates will not be as high at their peak, but will decline less steeply with age.

2.3 Equilibrium sorting in the marriage market

Marriage market equilibria can be dramatically affected by imbalances in the populations of men and women. In a recent popular book, Birger (2015) provides numerous examples of how imbalanced sex ratios can have outsized impacts on dating culture and marriage markets in settings ranging from cities to college campuses to religious communities.[25] These examples illustrate that supply and demand quite literally rule the marriage market, and so it is critical that any empirical study of marriage account for such equilibrium sorting. Equilibrium sorting under a relative shortage of men can also generate the marriage and remarriage patterns observed in Figure 2.1. If more women than men are seeking to marry, then men will enjoy greater success in the marriage market, by the logic of supply and demand. Such a situation could arise because of factors that remove men from the marriage market, such as migration, mortality, and incarceration. In China, the reverse has occurred — the One Child Policy has led to an extreme shortage of girls in some areas (due to a cultural preference for boys and the widespread practice of selective abortions). Bride prices — customary gifts of money and property from the groom's family to the newly-wed couple — have skyrocketed due to the sex ratio imbalance in the marriage market.[26]

A number of papers have found large effects of marriage market sex ratios on female marital outcomes.[27] Guttentag and Secord (1983) argues that the post-WWII Baby Boom

[25]The book discusses at length the difficulties of dating for women (and the advantages for men) in New York City and contrasts it with West Coast cities such as Seattle and San Jose to illustrate how sex ratio imbalances affect peoples' dating experiences. Another interesting example covered is the dating cultures of college campuses with large sex ratio imbalances, at both ends of the spectrum. Not surprisingly, long-term monogamous relationships are the norm when women are scarce, and short-term encounters are more typical when men are scarce.

[26]For example, "A distorted sex ratio is playing havoc with marriage in China" (2017) reports that in the province of Shandong, which lies between Beijing and Shanghai, the sex ratio reached 123:100 in 2010. Local bride prices have increased 100-fold in some villages there over the past ten years.

[27]For example, Angrist (2002) studies marriage among immigrant populations, exploiting the fact that most marriages were within ethnic groups. Another example is Abramitzky, Delavande, and Vasconcelos

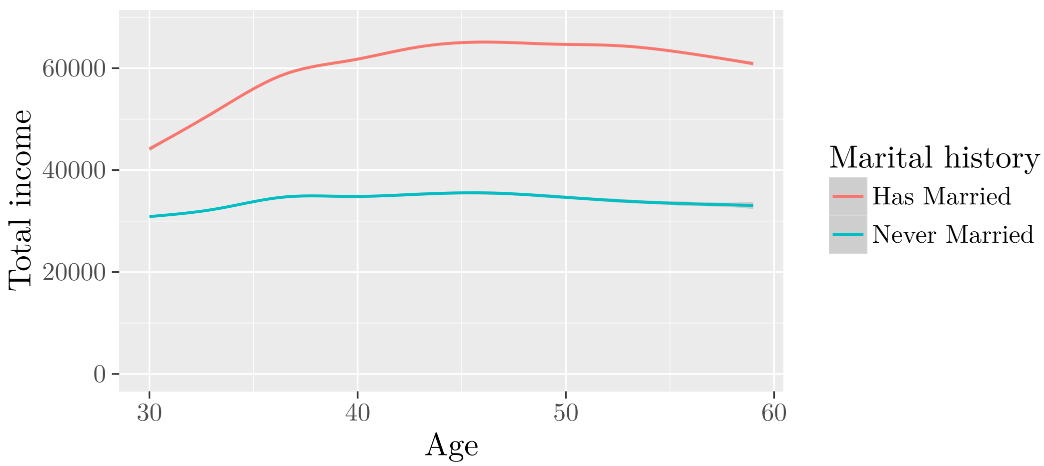
caused a "marriage squeeze"[28] for women in that birth cohort, which in turn was a fundamental driver of the many social upheavals of the 1960s, as women were driven by poor marriage prospects to pursue independence from marriage. Bronson and Mazzocco (2018) use an overlapping-generations model of marriage search to capture this cohort-size effect and finds that it can explain most of the variation in marriage rates since the 1930s. A similar paper is Rıos-Rull, Seitz, and Tanaka (2016), which also estimates an overlapping-generations model of marriage search to show that changes in demographics, via their effect on sex ratios, can almost perfectly explain the observed changes in marriage rates from 1870 to 1950.

Moreover, even if enough men are physically present, many are considered unmarriageable due to chronic unemployment, mental illness, substance abuse, crime, etc. This is the hypothesis advanced by Wilson (1987), who argued that the low marriage rates of black women are attributable to the poor labor market prospects of black men. By the same token, Lundberg and Pollak (2007) note in their review of American marriage trends that "the deteriorating market prospects of less-educated men during the 1980s and 1990s may have played a role in increasing nonmarital childbearing". Autor, Dorn, Hanson, et al. (2017) apply this logic to study how male employment prospects affect female marriage and childbearing patterns. They find that trade shocks which negatively impacted male employment in manufacturing decreased the prevalence of marriage among women; conversely, negative shocks to female employment increased the prevalence of marriage among women.

In this case, the least marriageable men may not be able to marry at all, as women will opt to marry older divorced men in their stead. By virtue of the fact that they were considered suitable for marriage in the first place, divorced men will tend not to be among the least desirable men. Thus, a group of the least marriageable men will not marry at all, while other men will marry multiple times. This "serial polygyny" equilibrium generates

(2011), which uses local variation in mortality from WWI in France to identify the effect of sex ratios on marriage outcomes.

[28]When combined with a marital age gap, sudden changes in cohort size result in sex ratio imbalances. An increase in cohort size leaves the women in that cohort facing a relative shortage of men in the cohorts a few years older than them. Similarly, a decrease in cohort size leaves the men in that cohort facing a relative shortage of women in the cohorts a few years younger than them.

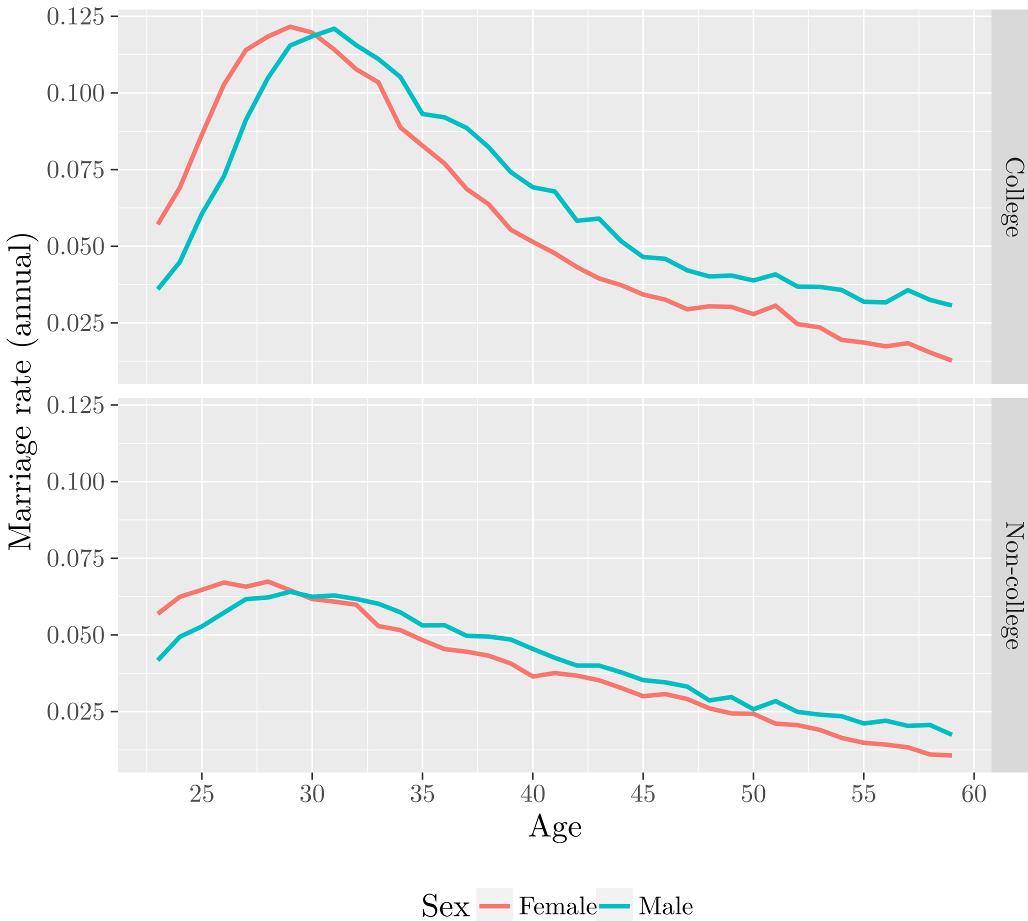


Notes: Men are categorized by whether they have ever been married. ACS 2010–2014.

Figure 2.5: Male income by ever-married status

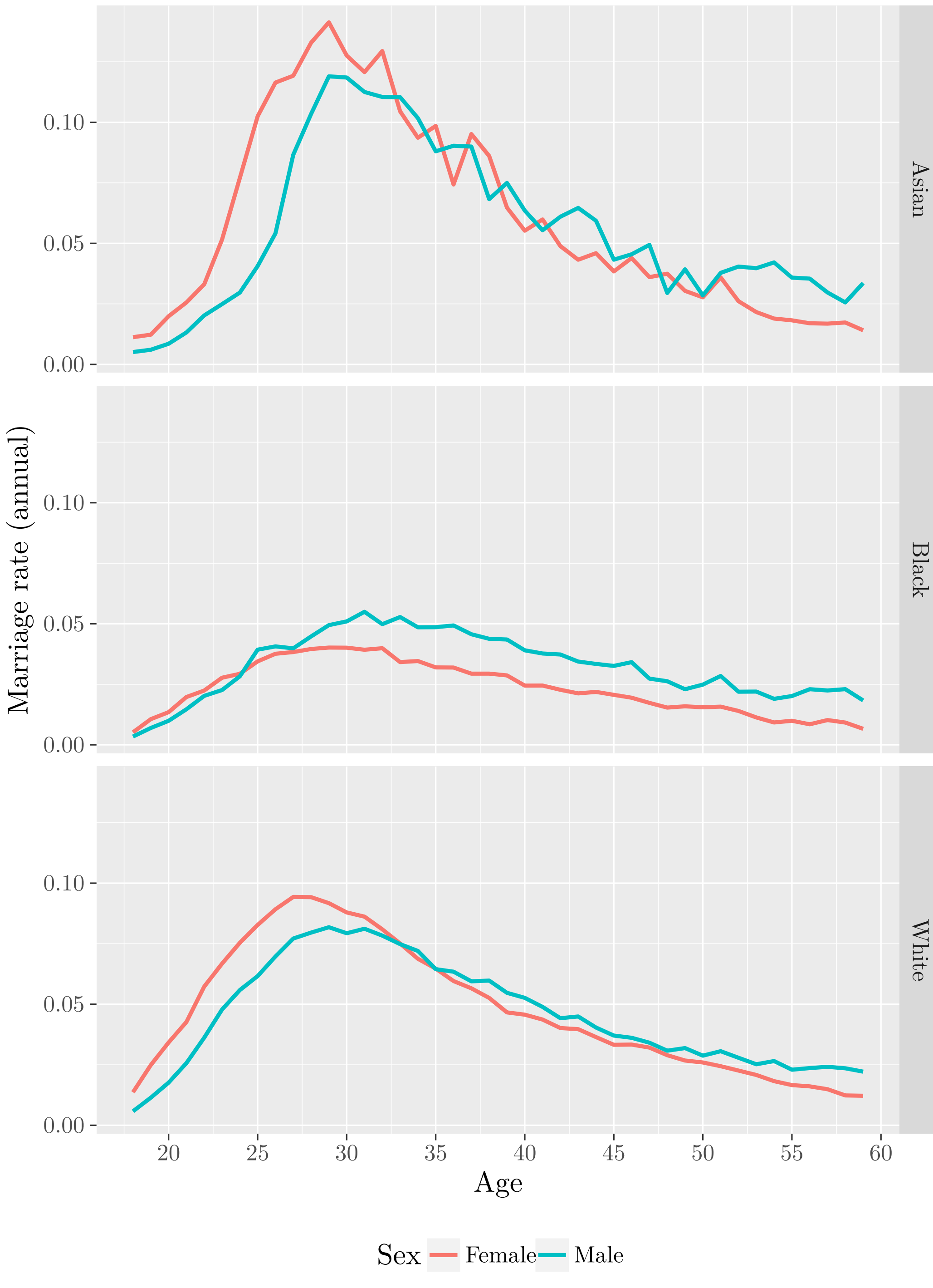
both a lower first marriage rate and a higher remarriage rate for males, as observed in the data. Figure 2.5 shows that, by middle age, married men earn almost twice as much as never married men. Of course, some of this difference in earnings is due to endogenous choices such as working harder to support a family. However, it is hard to believe that selection into marriage does not play a substantial role in generating this difference. As income is an important factor in a man's desirability as a husband, women will consider marrying higher-earning divorced men rather than lower-earning never-married men.

Sex ratio imbalances can also be concentrated within groups when members prefer to marry within the group. An unusually clear example of this is provided by minority religious groups. Birger (2015) notes that the sex ratio among Mormons in Utah is 100:150, despite the fact that the state of Utah has more men than women overall. The reason for this is that Mormon men have been leaving the church at far higher rates than women over the past three decades. This has made dating very competitive for Mormon women, who face a race against time to get married before their age makes them uncompetitive in the marriage market. More generally, sex ratio imbalances within groups affect marriage rates in the expected way. Figures 2.6 and 2.7 show overall marriage rates by college attainment and race. Figure F.1 in Appendix F shows the overall marriage rates by urban status. In each



Notes: The marriage rate is the proportion of all singles who enter into marriage during the survey year. ACS 2008–2016.

Figure 2.6: Marriage rates by sex and education



Notes: The marriage rate is the proportion of all singles who enter into marriage during the survey year. ACS 2008–2014.

Figure 2.7: Marriage rates by sex and race

case, women have relatively higher marriage rates at younger ages and men have relatively higher marriage rates at later ages. But the variation between groups is highly suggestive of the effects of a sex ratio imbalance — the marriage rate gap between men and women at later ages is far larger among groups in which there are relative shortages of men: college graduates, blacks, and urban populations.[29]

My equilibrium model of the marriage market takes account of the populations of men and women by age as well as by education and race. This allows it to capture the complexities of matching across these dimensions. Education acts as a proxy for income and marriageability more generally. I also estimate my model at the local level in order to capture the relative suppy and demand conditions of different marriage markets, arising from heterogeneity of their populations.

2.4 Other hypotheses

Finally, it could be that women have close substitutes for husbands (and men do not have close substitutes for wives). Single mothers often qualify for welfare benefits. In the case of divorce, a woman may receive alimony and/or child support payments from her ex-husband. These income sources relieve the financial burden that would otherwise increase the incentive for single women to marry.

Another anecdotal story is that older women simply have less to gain from marriage. However, this story is confounded by marital opportunities — if the quality of men willing to marry them is too low, women will forgo these low-value marriages and will correctly state that it is because they do not have much to gain from such marriages.

[29] As noted above, women complete college at much higher rates than men, a gap that has widened since women surpassed men in the 1980s. This is directly visible in the population supplies of people with and without college degrees (compare Figures F.2 and F.3 for non-minorities, or Figures F.4 and F.5 for minorities). The sex ratio imbalances within educational classes are especially pronounced for younger generations: there are far more college-educated women than men in these cities, and far more men than women among those without a college degree. The same pattern holds for race, with there being relatively fewer men among minorities. In combination, this results in quite extreme sex ratio imbalances for college-educated minorities (Figure F.4).

Chapter 3

Model

The marriage market is represented by an equilibrium search and matching model which features overlapping generations and an aging process. The model builds on the equilibrium search-bargaining model of Shimer and Smith (2000) and the extensions of Goussé, Jacquemet, and Robin (2017b) and Goussé (2014). In particular, the model bears greatest resemblance to that of Goussé (2014), and as the notation imposed by the aging process can be cumbersome, I refer the reader to Appendix D for a brief review of that model.

These models rely on the assumption steady state equilibrium for identification as well as solution. In particular, this simplifies the problem of agent expectations about future states of the marriage market. As populations of singles are unchanging, agents need not forecast the future to form expectations.

3.1 The marriage market

Consider a single marriage market. There are continuua of men and women who search randomly for matches in continuous time. People discount the future at a common rate r. Each person is fully characterized by their age and a type vector. Throughout, I adhere to the following notation for functions of individual and couple ages and type vectors:

1. Sex (m, f): subscript

2. Age (e.g., a, b): superscript

3. Type vector (e.g., x, y): function arguments

In the text, I sometimes use tuples to refer to age-type combinations, e.g., (a, x) men or (b, y) women. As the sexes are symmetric in the model, I present the equations for males and omit the female cases for the sake of brevity.

3.1.1 Populations

Denote the male and female population measures by $\ell_m^a(x), \ell_f^b(y)$, the singles measures by $u_m^a(x), u_f^b(y)$, and the measures of married couples by $m^{a,b}(x, y)$. Note that these are not probability distributions (i.e., they are not normalized). As such, the model allows for global sex ratio imbalances. Denote the total measures of males and females by $L_m := \sum_a \int \ell_m^a(x)dx$ and $L_f := \sum_b \int \ell_f^b(y)dy$. Similarly, denote the total measures of singles by $U_m := \sum_a \int u_m^a(x)dx$ and $U_f := \sum_b \int u_f^b(y)dy$. Men and women have different life expectancies, and die at age- and type-dependent Poisson rates $\psi_m^a(x), \psi_f^b(y)$.

The marriage market can be considered as a closed system with population inflows and outflows only from birth and death. It is also simple to allow for exogenous migration flows, though I do not model this at the individual level — people make decisions as if they will never migrate. Allowing for population-level migration flows helps the model to fit lifecycle migration patterns in the empirical application (e.g., young people migrating into and old people emigrating out of cities). Birth and migration are exogenous and the inflows of new people are denoted by $\gamma_m^a(x), \gamma_f^b(y)$. People migrate out of the marriage market at exogenous rates $\varphi_m^a(x), \varphi_f^b(y)$. Also denote couple-specific migration inflows by $\gamma^{a,b}(x, y)$ and the couple-specific outflow rate by $\varphi^{a,b}(x, y)$.

3.1.2 Stochastic aging

Although types are static, people proceed sequentially through discrete ages $1, \ldots, T$ by means of exogenous stochastic aging shocks. Stochastic aging is the most natural and tractable way to incorporate an aging process into the search model. This way, there is no need to keep track of time as a state variable, as aging shocks are treated as Poisson arrivals along with other stochastic events such as meetings and death.[30] An alternative approach is to use a deterministic continuous aging process, as in Coles and Francesconi (2011), but this makes the model much harder to solve.[31] I denote age separately from the type vector to make clear that it is a dynamic state variable.

Let ρ be the Poisson arrival rate of aging shocks. Each arrival increments a person's age by 1. In the terminal age category T, aging stops (but everyone dies eventually because of arrivals of death shocks). I assume that, for a married couple, aging shocks are perfectly correlated, so that the pair ages together in lock-step.

3.1.3 Meeting technology

Single men and women meet one another at random through a frictional search process. There is no search within marriage. Meetings are modeled as stochastic Poisson arrivals, and meeting rates may differ based on types. I follow the approach of Shin (2014), whose model features race-specific meeting rates, which she uses to separately identify marital preferences and dating opportunities with data on interracial marriage patterns. Here, meeting rates are allowed to vary by the ages or types of each pair. This reflects the fact that people may be more likely to meet others of similar age, race, and education.

The flow rate of meetings between single (a, x) men and single (b, y) women is given by

$$M^{a,b}(x,y) := \xi^{a,b}(x,y) M\left(U_m, U_f\right) \frac{u_m^a(x)}{U_m} \frac{u_f^b(y)}{U_f},$$

[30]Poisson events are a convenient distribution for events in stationary continuous-time models because inter-arrival durations are exponentially distributed. The exponential distribution is the only memoryless continuous distribution, which means that the distribution of waiting times is independent of how much time has already elapsed. Without memorylessness, it would be necessary for agents to keep track of time as a state variable.

[31]By using a constant rate of aging, this method also circumvents the need to track time as a state variable. See Appendix E for an overview of the model with a continuous aging process and the difficulties that arise in solving it.

where M is assumed to be a CRS[32] meeting technology. In the application, I use

$$M(U_m, U_f) := \sqrt{U_m U_f}.$$

Here, $M\left(U_m, U_f\right)$ can be interpreted as the overall rate of meetings in the marriage market, which is rescaled by $\xi^{a,b}(x,y)$ to reflect differing efficiencies in generating meetings between types. These meetings are distributed among the different pairs of types according to their share of the singles population: $\frac{u_m^a(x)}{U_m}\frac{u_f^b(y)}{U_f}$.

Define the meeting function $\lambda^{a,b}(x,y)$ as the flow of meetings per unit of time divided by the total number of potential meetings between (a,x) men and (b,y) women. In other words, $\lambda^{a,b}(x,y)$ is the meeting rate: the fraction of potential meetings that are realized per unit of time. Formally,

$$\lambda^{a,b}(x,y) = \lambda^{a,b}\left(x,y;U_m,U_f\right) := \frac{M^{a,b}(x,y)}{u_m^a(x)u_f^b(y)} = \xi^{a,b}(x,y)\frac{M(U_m,U_f)}{U_m U_f}.$$

Thus, the flow of meetings between (a,x) men and (b,y) women can be written as

$$M^{a,b}(x,y) = \lambda^{a,b}(x,y)u_m^a(x)u_f^b(y).$$

For an (a,x) man, the arrival rate of meetings with (b,y) women is

$$\frac{M^{a,b}(x,y)}{u_m^a(x)} = \lambda^{a,b}(x,y)u_f^b(y),$$

and so the arrival rate of meetings with any woman is

$$\frac{1}{u_m^a(x)}\sum_b \int M^{a,b}(x,y)dy = \sum_b \int \lambda^{a,b}(x,y)u_f^b(y)dy.$$

The fact that M is a CRS technology means that, for an unmarried person, the rate at which meetings arrive is invariant to the total number of singles in the marriage market,

[32] CRS meeting technology is also used in Goussé, Jacquemet, and Robin (2017b).

i.e., doubling the number of single men and women has no effect on the meeting rates experienced by individuals. Rewriting the expression for $\lambda^{a,b}(x,y)u_f^b(y)$, it is easy to see that it is homogenous of degree 0 in the singles population: $\forall t > 0$,

$$\begin{aligned}
\xi^{a,b}(x,y)\frac{M((tU_m),(tU_f))}{(tU_m)(tU_f)}(tu_f^b(y)) &= \xi^{a,b}(x,y)\frac{tM(U_m,U_f)}{t^2U_mU_f}tu_f^b(y) \\
&= \xi^{a,b}(x,y)\frac{M(U_m,U_f)}{U_mU_f}u_f^b(y) \\
&\equiv \lambda^{a,b}(x,y)u_f^b(y).
\end{aligned}$$

Clearly, this also holds when integrating over all ages and types of women, which gives the total rate of meetings.

Intuitively, this means that singles are able to meet one another at the same rate, regardless of how many singles there are in the market. This makes sense for the obvious reason that people are time-constrained in how many other people they can meet, and this constraint is certainly binding for almost everyone living in a moderately populated city. In other words, the local population size is not the limiting factor of the number of meetings a person can realize.

With CRS meeting technology, sex ratios alone determine the meeting rate experienced by individuals. However, for meetings with singles of a specific type, the relative population share of that type does matter. With the total number of singles held fixed, an (a,x) man meets (b,y) women at rate $\lambda^{a,b}(x,y)u_f^b(y)$, i.e., at a rate proportional to the share of (b,y) women in the singles population. So, even though the matching technology is CRS with respect to the overall population of singles, it is quadratic with respect to the relative shares of singles by type.

As such, there is a matching externality: when a (b,y) woman gets married, the relative share of (b,y) singles, $u_f^b(y)/U_f$, shrinks, reducing the rate at which men meet them (though this is offset by more meetings with other types whose relative shares necessarily increased). In particular, this can give rise to a strategic complementarity in marriage timing. If people prefer to match with someone of similar age, then there is an incentive to coordinate on marriage timing. Those who postpone marriage for too long will find that meetings with other singles their age are hard to come by, as most of their meetings will be from the more plentiful group of younger singles.

3.1.4 Marriage, divorce, and bargaining

Upon meeting, singles decide whether to marry or to keep searching. The gains from marriage are represented by a systematic component based on the pair's observed types, and an idiosyncratic component representing unobserved heterogeneity in match quality. I normalize the flow value of singlehood to 0 without loss of generality, as marriage and divorce decisions only depend on the gains from marriage relative to singlehood. The systematic component is modeled as a marital production function, $f^{a,b}(x, y)$, which is the flow utility generated by a marriage between an (a, x) man and a (b, y) woman. The idiosyncratic component is modeled as a random, match-specific love shock that is drawn upon meeting, before the marriage decision:

$$z \overset{iid}{\sim} G.$$

There is no learning or uncertainty; when a pair meets, they observe their love shock and make an instantaneous marriage decision. This effectively means that the process of courtship is instantaneous.

If they marry, the resulting marital flow output is $f^{a,b}(x, y) + z$. Utility is perfectly transferable within marriage, and the marital flow output is divided according to Nash Bargaining, with the outside options of being single and searching as threatpoints. The assumption of perfectly transferable utility is crucial for identification of the model, as will become evident below. This is a reasonable assumption, as there are manifold ways for spouses to compensate one another within marriage.[33]

[33] Money is an obvious way for spouses to compensate one another; housework is another. Beyond this, there is plenty of scope for bargaining over major joint decisions such as: family planning (how many children to have and when), parenting (how to raise the children), where to live (close to whose family, close to which labor market), career decisions (whether one spouse should make career sacrifices to spend more time raising children, when to retire), etc.

The division of the marital flow output between husband and wife is denoted by

$$f^{a,b}(x,y) + z = t_m^{a,b}(x \mid y,z) + t_f^{a,b}(y \mid x,z). \tag{3.1.1}$$

Married people also make separation choices. I adopt the endogenous divorce process developed by Goussé (2014). Over the course of a marriage, love shocks $z \sim G$ are renewed at Poisson rate δ, triggering a renegotiation of the intramarital transfers. If the updated love shock is insufficient to sustain the marriage, the couple divorces. This assumption of no commitment in marriage is largely irrelevant in the context of transferable utility. In the model, separation decisions are always mutual, i.e., a couple will choose to divorce if and only if they would both be better off apart. By contrast, in a model with non-transferable utility, the assumption of no commitment would mean that there could be unilateral divorces in which one spouse is made worse off.

It is worth noting that the love shocks are i.i.d. (independent and identically distributed), although a more realistic assumption would be that they are serially correlated. This would be the case for many sources of unobserved heterogeneity which are highly persistent (e.g., personality, values, habits, etc). With i.i.d. love shocks, a high realization of z may not be sufficient to warrant a marriage, as the shock is transitory and renewals will revert to the mean. With serially correlated love shocks, a high realization of z would imply a more durable marriage. Hence, the model undervalues the persistence of the idiosyncratic gains from marriage (z) and will therefore underestimate the durability of marriages that have low systematic gains ($f^{a,b}(x,y)$) based on observables. However, allowing for serial correlation would require keeping track of z as a state variable, which greatly complicates the model and is beyond the scope of this paper. Rıos-Rull, Seitz, and Tanaka (2016) use a simple two-state Markov process, but they estimate the model by simulation. A more general approach would be a Markov process with a continuous state-space, e.g., an AR(1) process (auto-regressive process of order 1) as commonly used in time-series modeling.

At every age transition, the marital output $f^{a,b}(x,y)$ and outside options change, and so the couple renegotiates the division of the marital surplus via Nash Bargaining. To maintain tractability, I assume that marriages can only be dissolved upon arrival of a new love shock, and not at an age transition. Due to aging, it is possible that a couple's marital surplus becomes negative (it must start out positive) so that they would both be better off as singles. However, I impose the restriction that they must remain married in this unhappy state until an updated love shock separates them.

3.2 Steady state population measures

Given birth and migration inflows, the steady state population measures in the marriage market are derived from the condition that population inflows equal population outflows for people of each sex, age, and type:

$$(\rho + \varphi^a_m(x) + \psi^a_m(x))\ell^a_m(x) = \gamma^a_m(x) + \rho\ell^{a-1}_m(x). \tag{3.2.1}$$

The population sizes can then be solved forward from $a = 1$, using the boundary condition $\ell^0_m \equiv 0$.

3.3 Strategies

Denote the discounted present value of an (a, x) man by $V^a_m(x)$ when single, and by $W^{a,b}_m(x \mid y, z)$ when married to a (b, y) woman with realized love shock z. Define the personal surplus as

$$S^{a,b}_m(x \mid y, z) := W^{a,b}_m(x \mid y, z) - V^a_m(x).$$

The value function for a single man of type x and age $a < T$ is then

$$(r + \rho + \psi^a_m(x))V^a_m(x) = \rho V^{a+1}_m(x) + \sum_{b=1}^{T} \iint \lambda^{a,b}(x, y) \max\left\{S^{a,b}_m(x \mid y, z), 0\right\} u^b_f(y)dG(z)dy. \tag{3.3.1}$$

This expression takes into account both the continuation value from aging as well as the expected value of a potential marriage.

When a man of type x and age $a < T$ is married to a type y woman of age $b < T$, the value function is

$$(r + \rho + \psi_m^a(x) + \psi_f^b(y))W_m^{a,b}(x \mid y, z) = t_m^{a,b}(x \mid y, z) + \rho W_m^{a+1,b+1}(x \mid y, z) \\ + \psi_f^b(y)V_m^a(x) - \delta \left(S_m^{a,b}(x \mid y, z) - \int \max\{S_m^{a,b}(x \mid y, z), 0\} dG(z) \right). \quad (3.3.2)$$

The derivations for these value functions are provided in Appendix C. Because aging stops once a person reaches age T, these terminal boundary cases are slightly different. I assume that aging shocks still arrive for a couple, but only affect the spouse that is younger than T. In other words, ages are truncated at T. For a single man of age T, there is no longer any continuation value from aging:

$$(r + \psi_m^T(x))V_m^T(x) = \sum_{b=1}^{T} \iint \lambda^{T,b}(x, y) \max \left\{ S_m^{T,b}(x \mid y, z), 0 \right\} u_f^b(y) dG(z) dy.$$

When a man of age T is married to a woman of age $b < T$, the value function is

$$(r + \rho + \psi_m^T(x) + \psi_f^b(y))W_m^{T,b}(x \mid y, z) = t_m^{T,b}(x \mid y, z) + \rho W_m^{T,b+1}(x \mid y, z) \\ + \psi_f^b(y)V_m^T(x) - \delta \left(S_m^{T,b}(x \mid y, z) - \int \max\{S_m^{T,b}(x \mid y, z), 0\} dG(z) \right).$$

Finally, when a couple are both of age T, the effect of aging shocks disappears:

$$(r + \psi_m^T(x) + \psi_f^T(y))W_m^{T,T}(x \mid y, z) = t_m^{T,T}(x \mid y, z) + \psi_f^T(y)V_m^T(x) \\ - \delta \left(S_m^{T,T}(x \mid y, z) - \int \max\{S_m^{T,T}(x \mid y, z), 0\} dG(z) \right).$$

Denote the total match surplus by

$$\mathcal{S}^{a,b}(x, y, z) := S_m^{a,b}(x \mid y, z) + S_f^{a,b}(y \mid x, z).$$

With Nash Bargaining over the match surplus, if women have bargaining power β, then the division of the surplus can be written as

$$\begin{cases} S_f^{a,b}(y \mid x, z) & = \beta S^{a,b}(x, y, z) \\ S_m^{a,b}(x \mid y, z) & = (1-\beta) S^{a,b}(x, y, z). \end{cases}$$

Because utility is transferable, the marriage decision is always mutual: a pair marries if and only if $\mathcal{S}^{a,b}(x, y, z) > 0$. Using Equation (3.1.1), adding up the value functions for a married couple yields an expression for the total marital surplus. As such, the model can be solved without any need to keep track of the individual transfers t_m, t_f (the personal gains from marriage). For $a, b < T$,

$$\begin{aligned} (r + \rho + \delta + \psi_m^a(x) + \psi_f^b(y))\mathcal{S}^{a,b}(x, y, z) &= z + f^{a,b}(x, y) + \rho \mathcal{S}^{a+1,b+1}(x, y, z) \\ &+ \delta \int \max\left\{\mathcal{S}^{a,b}(x, y, z), 0\right\} dG(z) + \rho(V_m^{a+1}(x) + V_f^{b+1}(y)) \\ &- (r + \rho + \psi_m^a(x))V_m^a(x) - (r + \rho + \psi_f^b(y))V_f^b(y). \quad (3.3.3) \end{aligned}$$

When one spouse reaches the terminal age T,

$$\begin{aligned} (r + \rho + \delta + \psi_m^a(x) + \psi_f^T(y))\mathcal{S}^{a,T}(x, y, z) &= z + f^{a,T}(x, y) + \rho \mathcal{S}^{a+1,T}(x, y, z) \\ &+ \delta \int \max\left\{\mathcal{S}^{a,T}(x, y, z), 0\right\} dG(z) + \rho V_m^{a+1}(x) \\ &- (r + \rho + \psi_m^a(x))V_m^a(x) - (r + \psi_f^T(y))V_f^T(y). \end{aligned}$$

Finally, when both spouses reach the terminal age T, the aging process ceases,

$$\begin{aligned} (r + \delta + \psi_m^T(x) + \psi_f^T(y))\mathcal{S}^{T,T}(x, y, z) &= z + f^{T,T}(x, y) + \delta \int \max\left\{\mathcal{S}^{T,T}(x, y, z), 0\right\} dG(z) \\ &- (r + \psi_m^T(x))V_m^T(x) - (r + \psi_f^T(y))V_f^T(y). \quad (3.3.4) \end{aligned}$$

3.4 Matching Equilibrium

A *matching equilibrium* is a partial equilibrium concept which describes the best-response (i.e., Nash) matching strategies. Taking the measures of singles as exogenously given, a matching equilibrium is the set of acceptance rules that govern matching decisions. As matching decisions are always mutual (because of transferable utility), there is no need to keep track of separate strategies for men and women, and strategies can be reduced to the pairwise decision rules, $\mathcal{S}^{a,b}(x,y,z) > 0$. Accounting for the randomness introduced by the match-specific love shocks z, the matching outcome is captured by $\alpha^{a,b}(x,y)$, the probability that an (a,x) man and a (b,y) woman will marry conditional upon meeting.

A match is formed if and only if $\mathcal{S}^{a,b}(x,y,z) > 0$. In order to derive the conditional match probability function $\alpha^{a,b}(x,y)$, I first show that $\mathcal{S}^{a,b}(x,y,z)$ is linear in z. Starting from the terminal states, Equation (3.3.4) can be written as

$$(r + \delta + \psi_m^T(x) + \psi_f^T(y))\mathcal{S}^{T,T}(x,y,z) = z + s^{T,T}(x,y)$$

for some function $s^{T,T}$. In other words, the average marital surplus can be decomposed into a sum of an idiosyncratic component and a systematic component. This gives

$$\alpha^{T,T}(x,y) = \mathbb{P}\left[\mathcal{S}^{T,T}(x,y,z) > 0\right] = \mathbb{P}\left[z > -s^{T,T}(x,y)\right] = 1 - G(-s^{T,T}(x,y)).$$

This linear decomposition of $\mathcal{S}$ can also be done for younger ages by recursively working backward from this boundary case. First, I introduce some notation to keep track of the discount factors. Define

$$d^{a,b}(x,y) := \begin{cases} \left(r + \rho + \delta + \psi_m^a(x) + \psi_f^b(y)\right)^{-1}, \ a < T \text{ or } b < T \\ \left(r + \delta + \psi_m^T(x) + \psi_f^T(y)\right)^{-1}, \ a = b = T. \end{cases}$$

Define $c^{T,T}(x,y) := 1 + \rho d^{T,T}(x,y)$, and for $a,b < T$, define $c^{a,b}(x,y)$ recursively as follows:

$$c^{a,b}(x,y) := \begin{cases} 1 + \rho d^{T,T}(x,y),\ a = b = T \\ 1 + \rho d^{a,b}(x,y) c^{a+1,b+1}(x,y),\ a,b < T \\ 1 + \rho d^{a,T}(x,y) c^{a+1,T}(x,y),\ a < T, b = T \\ 1 + \rho d^{T,b}(x,y) c^{T,b+1}(x,y),\ a = T, b < T. \end{cases}$$

Using $\bar{s}^{a,b}(x,y)$ to denote all of the terms of Equation (3.3.3) that do not depend on z,

$$\frac{\mathcal{S}^{a,b}(x,y,z)}{d^{a,b}(x,y)} = z + \rho \mathcal{S}^{a+1,b+1}(x,y,z) + \bar{s}^{a,b}(x,y).$$

Working back recursively from the terminal states gives the following result:

$$\mathcal{S}^{a,b}(x,y,z) = d^{a,b}(x,y)\left(z \cdot c^{a+1,b+1}(x,y) + s^{a,b}(x,y)\right), \tag{3.4.1}$$
$$\text{where } s^{a,b}(x,y) := \bar{s}^{a,b}(x,y) + \rho d^{a+1,b+1}(x,y) s^{a+1,b+1}(x,y).$$

The other boundary cases are similar, as aging shocks continue to arrive so long as one spouse is younger than T. Ages are truncated at T, but the younger spouse continues to age as usual. For example,

$$\mathcal{S}^{T,b}(x,y,z) = d^{T,b}(x,y)\left(z \cdot c^{T,b+1}(x,y) + s^{T,b}(x,y)\right).$$

Thus, $\mathcal{S}^{a,b}$ is linear in z and can be written in terms of $s^{a,b}$. This yields a solution for the conditional matching probability function, $\alpha^{a,b}(x,y)$:

$$\begin{aligned} \alpha^{a,b}(x,y) :&= \mathbb{P}\left[\mathcal{S}^{a,b}(x,y,z) > 0\right] \\ &= \mathbb{P}\left[z > -\frac{s^{a,b}(x,y)}{c^{a+1,b+1}(x,y)}\right] \\ &= 1 - G\left(-\frac{s^{a,b}(x,y)}{c^{a+1,b+1}(x,y)}\right). \end{aligned} \tag{3.4.2}$$

3.5 Market Equilibrium

To close the model, I invoke the steady state assumption to require that marriage flows are balanced: outflows equal inflows for every kind of marriage. The resulting steady state is called a *market equilibrium.* Whereas a matching equilibrium is a partial equilibrium concept that treats the state of the marriage market as exogenous, a market equilibrium is a general equilibrium concept.

Formally, $\left\{u_m^a, u_f^b, m^{a,b}\right\}_{a,b=1}^T$ solves the steady state flow condition: outflows equal inflows for every kind of marriage,

$$\begin{aligned}\left(\rho + \varphi^{a,b}(x,y) + \psi_m^a(x) + \psi_f^b(y) + \delta(1-\alpha^{a,b}(x,y))\right) m^{a,b}(x,y) = \\ \lambda^{a,b}(x,y)u_m^a(x)u_f^b(y)\alpha^{a,b}(x,y) + \rho m^{a-1,b-1}(x,y) + \gamma^{a,b}(x,y).\end{aligned} \tag{3.5.1}$$

The left hand side of this equation represents all sources of outflows from a particular marriage state: aging, migration, death, and divorce. The right hand side represents all sources of inflows into a particular marriage state: marriages, aging, and exogenous inflows. Here, the matching probability function $\alpha^{a,b}(x,y)$ is the endogenous outcome of the matching equilibrium, which is determined by $\left\{u_m^a, u_f^b, m^{a,b}\right\}_{a,b=1}^T$. Put another way, a market equilibrium requires that the marriage flows implied by $\alpha^{a,b}(x,y)$ are consistent with the steady state measures of singles and marriages.

This equation is slightly different for the ages (T,T) boundary case, as aging stops $(\rho = 0)$:

$$\begin{aligned}\left(\varphi^{T,T}(x,y) + \psi_m^T(x) + \psi_f^T(y) + \delta(1-\alpha^{T,T}(x,y))\right) m^{T,T}(x,y) = \\ \lambda^{T,T}(x,y)u_m^T(x)u_f^T(y)\alpha^{T,T}(x,y) \\ + \rho\left(m^{T-1,T-1}(x,y) + m^{T,T-1}(x,y) + m^{T-1,T}(x,y)\right) + \gamma^{T,T}(x,y).\end{aligned}$$

For age group 1, the boundary conditions are $m^{a,0}(x,y) = m^{0,b}(x,y) \equiv 0$. Entry of already-married couples into the marriage market via aging is represented by the inflow terms $\gamma^{a,0}(x,y), \gamma^{0,b}(x,y)$.

I outline a nested fixed-point algorithm for numerically solving a market equilibrium in Appendix B.

Chapter 4

Estimation

4.1 Data and calibration

The *American Community Survey*[34] provides rich cross-sectional microdata on an annual basis, from which stocks of singles and married couples, as well as flows of marriages and divorces can be computed at the MSA level. I pool the ACS 1% samples for the 2008–2016 survey years in order to obtain a sufficient sample size for estimating the model on MSA-level marriage markets. I restrict the analysis to the 20 largest MSAs (see Table F.1 in Appendix F for the full list).

I estimate two different model specifications. In the first specification, agents are characterized solely by their age. In the second specification, agents are additionally characterized by their education (whether they have a college degree or not), and race (whether they are a racial minority or not). I define the non-minority racial class as all whites and Asians, as the marriage patterns of Asians are much more similar to those of whites.[35]

The model uses a stochastic aging process, with age increasing over a finite set of age categories. As such, initial and terminal ages must be specified. For the first specification, I set the initial age at 19 with the 18 year-old population providing the inflow into the initial age group in the marriage market. For the second specification, the endogeneity of

[34] Data provided by IPUMS, Ruggles et al. (2017)

[35] Hispanics are also catergorized as racial minorities, though hispanic denotes an ethnic, not a racial group. For the precise race categorization used, see Section A.1 in Appendix A

education choices is a major concern, so I set the initial age to 26, at which point college completion has been determined for most people. In both specifications, the terminal age is set to 65, and everyone from age 66–79 is included in this group — this way, the model takes account of the marriage market equilibrium for older people, which can affect marriage market conditions for younger groups. I use actual calendar-year ages (i.e., 1-year age groups) which gives 47 age groups (19–65) for the first specification and 40 age groups (26–65) for the second specification. This way, the stochastic aging process in the model is a reasonable approximation to actual deterministic and continuous aging.

The (annual) discount rate is calibrated to $r = 0.04$.[36] I assume that men and women have equal bargaining power in marriage: $\beta = 0.5$. As durations between arrivals are exponentially distributed, the arrival rate parameters are simply the reciprocals of their respective mean interarrival durations.[37] Thus, the arrival rate of aging shocks is set to $\rho = 1$, as each age category in the model represents 1 year of calendar time.

Arrival rates of death shocks $\psi_m^a(x), \psi_f^b(y)$ are calibrated with mortality data from the CDC WONDER[38] database, using sex- and race-specific mortality rates. I transform annual mortality rates into arrival rate parameters using the CDF of the exponential distribution. Let D be a random variable for the time (in years) until death. If the annual mortality rate q is generated by an exponential distribution with rate parameter ψ, then

$$q = \Pr(D < 1) = 1 - \exp(-\psi) \iff \psi = -\log(1 - q).$$

Although the estimation is performed separately for each marriage market (MSA), I suppress the notation denoting the particular marriage market. Apart from arrival rate parameters, which are common to all markets, all other objects are specific to a particular marriage market. The marital production function $f^{a,b}(x, y)$, a model primitive, is estimated separately for each market and then the estimates are averaged together to arrive at a global

[36]Fixing the discount rate is a standard practice in the literature. For example, see Shin (2014) or Wong (2003a).

[37]If X is an exponentially distributed random variable with rate parameter λ, then $\mathbb{E}[X] = \frac{1}{\lambda}$.

[38]*Wide-ranging OnLine Data for Epidemiologic Research (WONDER)* (2017)

estimate.

4.2 Non-parametric smoothing of population measures

I first generate smoothed population measures for singles and couples from the raw counts in each MSA. This way, α is defined for every possible pair, and the estimates are less noisy. I similarly smooth the marriage and divorce flows, which are used to estimate the arrival rates in the model.

I smooth stocks and flows over the age dimensions by means of non-parametric local-polynomial regression. For individual measures such as ℓ, u, the smoothing is applied over age for each type and sex. I used local-linear regression with the cross-validated AIC bandwidth described in Li and Racine (2007) to smooth individual measures. Figure 2.3 shows the smoothed masses of individuals by age in the four largest MSAs.

For couples measures such as marriage stocks m, bivariate smoothing is applied over both ages for each pair of types. However, the cross-validated AIC bandwidth produced insufficiently smooth surfaces for couples measures. This is because the estimation is extremely sensitive to bumpiness in the age distributions, due to the necessity of calculating first differences along the age dimensions. As will be shown below, estimation of α requires taking differences of the form $m^{a,b} - m^{a-1,b-1}$. Estimation of the marital output f uses the estimated α and also requires another similar first difference. As such, the estimator of f tends to amplify small disturbances in the population measures.

To generate sufficiently smooth measures, I use a local-cubic regression along with a much wider bandwidth.[39] Wider bandwidths provide a higher degree of smoothing, but they also introduce bias (underfitting) because the estimator loses the ability to capture local details. Non-parametric regression is typically done with lower-order polynomials (most commonly local-linear) in order to avoid overfitting. But this also means that higher-order polynomials can be used to counteract underfitting resulting from a wide bandwidth. Considering that the measures I am smoothing are hump-shaped distributions, a cubic

[39] Goussé (2014) also describes requiring extra smoothing.

polynomial is sufficient to capture this shape over a larger bandwidth.

Because marriage is strongly assortative on age, the measures of couples have a simple hump shape that is oriented diagonally with respect to husband and wife ages. As such, it is best to orient the kernel diagonally as well, which is done by choosing an appropriate bandwidth matrix. Given a positive-definite and symmetric bandwidth matrix H, the kernel function is defined as

$$K_H(\mathbf{x}) = |H|^{-\frac{1}{2}} K\left(H^{-\frac{1}{2}}\mathbf{x}\right)$$

where $K(\mathbf{x}) = (2\pi)^{-1}\exp(-\frac{1}{2}\mathbf{x}'\mathbf{x})$ is a bivariate Standard Normal density and $\mathbf{x} = (a, b)'$ is a vector of the couples' ages. As such, the bandwidth matrix acts as a variance-covariance matrix for the kernel, allowing for it to be oriented in the direction of joint aging. Writing this matrix as

$$H = h_1 \times \begin{bmatrix} 1 & h_2 \\ h_2 & 1 \end{bmatrix},$$

h_1 is the variance and h_2 is the correlation of the kernel density. In conjunction with local-cubic regression, I set h_1 based on the model specification:

1. Age-only specification: $h_1 = 16$
2. Age, education, and race specification: $h_1 = 24$

With education and race as binary types, there are $2^2 \times 2^2 = 16$ possible couple types, each of which is smoothed over the age dimensions. This sharply exacerbates the data sparsity problem, especially for uncommon couple types, and so a wider bandwidth is warranted.

I set h_2 based on the object to be smoothed:

- Marriage stocks (m): $h_2 = 0.98$
- Marriage flows: $h_2 = 0.9$
- Migration flows (φ): $h_2 = 0.85$

These values were chosen to provide adequate smoothing without excessively underfitting the data. With ample data on marriage stocks, I use a large kernel correlation h_2, which places the local-regression weights mostly on couples with the same age gap (along the diagonal), even if they are much older or younger. With sparser data on flows, I reduce the kernel correlation h_2 to increase the degree of smoothing in the direction of couples with different age gaps, which helps to fill in values where there are no observations in the data.

Figure 4.1 shows the smoothed measures of married couples by age pair, $m^{a,b}$, in the four largest MSAs. Marriage is positive assortative on age, with the distribution closely following the $a = b$ line, with a slight bias toward husbands being older than wives.

4.3 Identification and estimation of arrival rates

The remaining arrival rate parameters, $\xi^{a,b}(x,y)$ and δ, are identified from data on marriage and divorce flows.[40] Denote the annual flows into and out of $((a,x),(b,y))$ marriages by $MF^{a,b}(x,y)$ and $DF^{a,b}(x,y)$, respectively. Then,

$$\begin{cases} MF^{a,b}(x,y) = \lambda^{a,b}(x,y)u_m^a(x)u_f^b(y)\alpha^{a,b}(x,y) \\ DF^{a,b}(x,y) = \delta(1-\alpha^{a,b}(x,y))m^{a,b}(x,y). \end{cases} \tag{4.3.1}$$

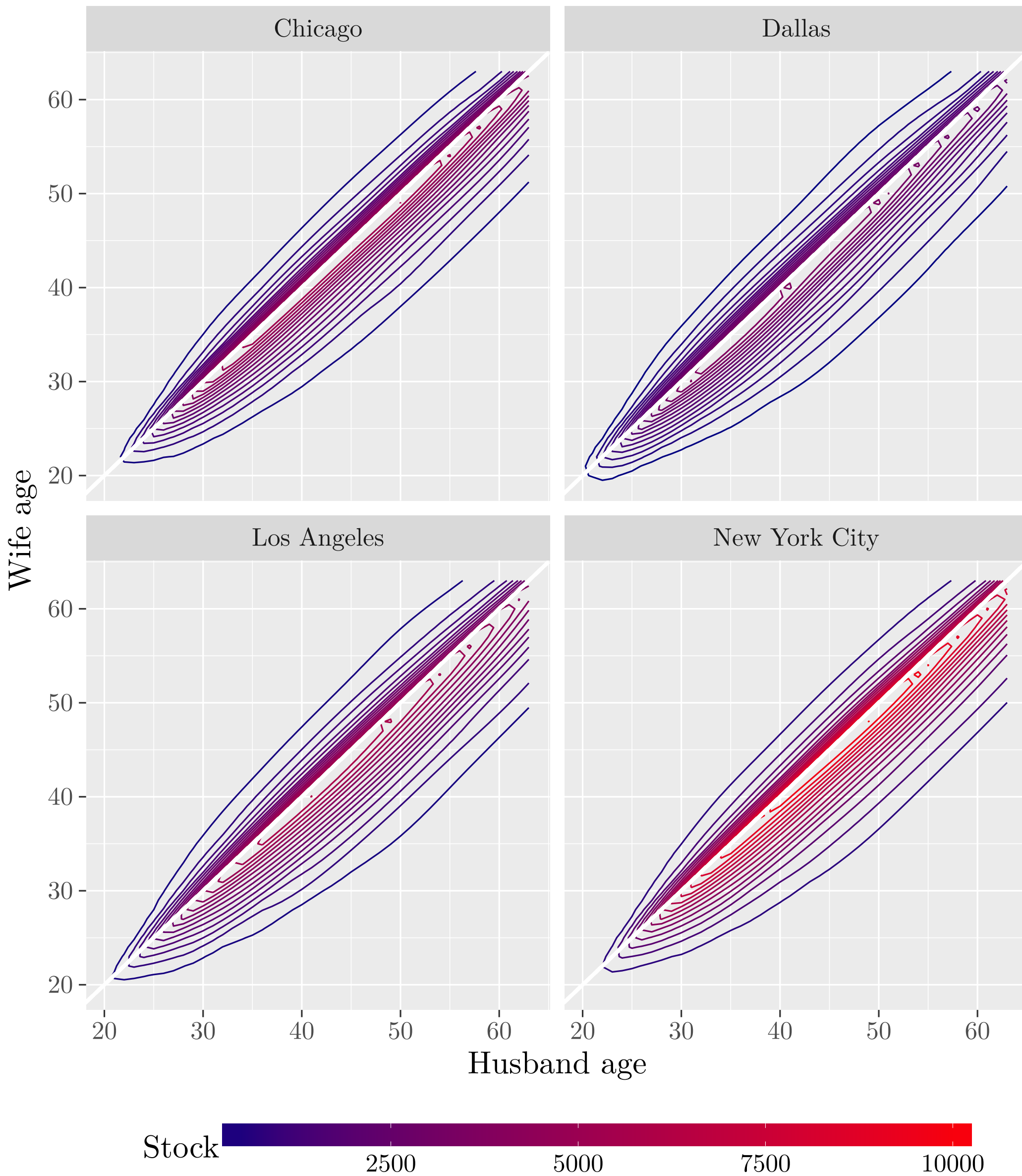
Also, $\alpha^{a,b}(x,y)$ can be solved by rearranging Equation (3.5.1):

$$\alpha^{a,b}(x,y) = \frac{(\rho+\varphi^{a,b}(x,y)+\psi_m^a(x)+\psi_f^b(y)+\delta)m^{a,b}(x,y)-\rho m^{a-1,b-1}(x,y)-\gamma^{a,b}(x,y)}{\lambda^{a,b}(x,y)u_m^a(x)u_f^b(y)+\delta m^{a,b}(x,y)}. \tag{4.3.2}$$

So, with data on marriage and divorce flows as well as population stocks for singles and couples, $\lambda^{a,b}(x,y) = \xi^{a,b}(x,y)\frac{M(U_m,U_f)}{U_m U_f}$ and δ are identified. The key assumption that allows meeting rates to be separately identified from marital output is that divorce decisions are no different from marriage decisions.[41] As love shocks are i.i.d. draws from the same distribution (regardless of whether it is the initial love shock or a renewal), α represents

[40] Goussé, Jacquemet, and Robin (2017b) derives a reduced-form equation in λ and δ that can be estimated by OLS regression using data on marriage and divorce flows by couple. See Section A.5 in Appendix A for details. Unfortunately, my data does not allow me to link divorces back to a couple and hence I cannot take this approach.

[41] This is the identification strategy developed in Shin (2014).



Notes: Counts are smoothed as described in Section 4.2. ACS 2008–2016.

Figure 4.1: Joint age distribution of married couples

strategies for both marriage and divorce. This way, the combination of both marriage and divorce flows separately identifies α and ξ. More intuitively, divorce flows can separately identify α because they are independent of meeting rates.

Although the ACS has rich data on annual divorce flows at the individual level, i.e., $DF^{a,b}(x,y)$, it is impossible to determine the ex-spouse associated with a given divorce. Thus, $DF^{a,b}(x,y)$ is not directly available in the data. Imputing divorce flows indirectly (from non-divorce and death flows) resulted in poor estimates due to the noise of the imputation.[42] Instead, I develop a GMM estimator for $\xi = \left[\xi^{a,b}(x,y)\right]$ and δ.

For moments, I use pairwise marriage flows and individual divorce flows (by age and type) for each MSA. The pairwise flows can be computed from Equation (4.3.1). The individual divorce flows are obtained by integrating out the opposite sex from the pairwise divorce flows:

$$DF_m^a(x) = \sum_{b=1}^{T} \int DF^{a,b}(x,y)dy.$$

I weight each pairwise marriage flow moment by

$$w^{a,b}(x,y) := \frac{\ell_m^a(x)\ell_f^b(y)}{L_m L_f},$$

the proportional representativeness of that pair of types (within a given MSA).[43] The individual divorce flow moments are weighted by the population proportion of each type, e.g., $w^a(x) := \ell_m^a(x)/L_m$.

To account for the fact that divorce flows are aggregated to the individual level by integrating over the opposite sex, the divorce flow moments must be scaled so that the squared moment deviations for marriage and divorce flows are on the same scale. Let d_f be the number of possible values for y (or the Lebesgue measure if y is continuous), so that $d_f T$ is the number (measure) of moments summed over in computing the divorce flow for

[42] See Section A.5 in Appendix A for details.

[43] This weighting scheme is also used by Goussé, Jacquemet, and Robin (2017b). Altonji and Segal (1996) provides evidence that GMM with optimal weight matrix is outperformed in practice by equal-weighted moments.

any given man. Then the divorce flows are rescaled in the natural way:

$$d_fT\left(\frac{DF_m^a(x;\xi,\delta)}{d_fT}-\frac{\widehat{DF}_m^a(x)}{d_fT}\right)^2=\frac{1}{d_fT}\left(DF_m^a(x;\xi,\delta)-\widehat{DF}_m^a(x)\right)^2.$$

Thus the estimator minimizes a Minimum Distance criterion function (denoting the MSA by M):

$$\begin{aligned}(\hat{\xi},\hat{\delta})=\arg\min_{\xi,\delta}\sum_M\Bigg\{&\sum_a^T\sum_b^T\iint w^{a,b}(x,y)\left(MF^{a,b}(x,y;\xi,\delta)-\widehat{MF}^{a,b}(x,y)\right)^2dxdy\\&+\frac{1}{2}\sum_{g\in\{m,f\}}\frac{1}{d_{-g}T}\sum_{k=1}^T\int w^k(v)\left(DF_g^k(v;\xi,\delta)-\widehat{DF}_g^k(v)\right)^2\Bigg\}dv.\end{aligned}\tag{4.3.3}$$

I compute bootstrap standard errors by resampling from the raw ACS data, followed by the smoothing and estimation steps. More details are provided in Section A.3 in Appendix A.

In practice, I do not estimate $\xi^{a,b}(x,y)$ pointwise (for every pair) because of the high dimensionality, though it is identified. Instead, I assume that the meeting rate is uniform across all pairs: $\xi^{a,b}(x,y)=\xi$. To the extent that meeting rates are non-uniform in actuality, this will bias the estimates of $\alpha^{a,b}(x,y)$ and ultimately $f^{a,b}(x,y)$. Most plausibly, meeting rates are endogenous and closely track the underlying marital productivity, as people will focus their search where the prospects are best. This will tend to exaggerate the estimates of $\alpha^{a,b}(x,y)$, which is decreasing in $\lambda^{a,b}(x,y)$, as can be seen from Equation (4.3.2). A more manageable approach to allowing meeting rates to vary would be to estimate a parametric functional form, but I do not explore that here.

4.4 Non-parametric identification and estimation

Once the arrival rate parameters are fixed, the other model objects can be non-parametrically identified, following the approach of Goussé, Jacquemet, and Robin (2017b). These are solved in the following order:

1. Conditional match probabilities: α

2. The systematic component of average marital surplus: s

3. Value functions of singlehood: V

4. Marital transfers: t; marital flow output; f.

In particular, there is no need to solve the model, as the primitives can be directly estimated from the equilibrium conditions of the model.

As the scale of $s^{a,b}(x,y)$ is not identified, the variance of the distribution of love shocks, G, can be normalized without loss of generality. I take G to be the Standard Normal distribution function. Then $s^{a,b}(x,y)$ is recovered by inverting Equation (3.4.2) (adjusting accordingly for the boundary cases):

$$s^{a,b}(x,y) = -c^{a+1,b+1}(x,y)G^{-1}(1-\alpha^{a,b}(x,y)).$$

To recover the marital output $f^{a,b}(x,y)$, it is necessary to first compute the singlehood value functions. This requires solving the integral in the value function Equation (3.3.1):

$$\begin{aligned}
&\int \max\left\{\mathcal{S}^{a,b}(x,y,z),0\right\} dG(z) \\
&\quad = d^{a,b}(x,y)\int_{-\infty}^{\infty} \max\{s^{a,b}(x,y)+c^{a+1,b+1}(x,y)z,0\}dG(z) \\
&\quad = d^{a,b}(x,y)c^{a+1,b+1}(x,y)\int_{-\frac{s^{a,b}(x,y)}{c^{a+1,b+1}(x,y)}}^{\infty}\left[\frac{s^{a,b}(x,y)}{c^{a+1,b+1}(x,y)}+z\right]dG(z) \\
&\quad = d^{a,b}(x,y)c^{a+1,b+1}(x,y)\alpha^{a,b}(x,y)\left(\frac{s^{a,b}(x,y)}{c^{a+1,b+1}(x,y)}+\mathbb{E}\left[z \mid z > -\frac{s^{a,b}(x,y)}{c^{a+1,b+1}(x,y)}\right]\right) \\
&\quad = d^{a,b}(x,y)c^{a+1,b+1}(x,y)\mu(\alpha^{a,b}(x,y)),
\end{aligned}$$

where $\mu(\zeta) := -G^{-1}(1-\zeta)+\mathbb{E}\left[z \mid z > G^{-1}(1-\zeta)\right]$.[44] Then the singlehood value functions

[44] A computationally-efficient formula for μ is given in Section A.4 in Appendix A.

can be recovered (adjusting accordingly for the boundary cases):

$$\begin{aligned}(r+\rho+\psi_m^a(x))V_m^a(x) &= \rho V_m^{a+1}(x) \\ &+ (1-\beta)\sum_{b=1}^{T}\int \lambda^{a,b}(x,y)d^{a,b}(x,y)c^{a+1,b+1}(x,y)\mu(\alpha^{a,b}(x,y))u_f^b(y)dy. \quad (4.4.1)\end{aligned}$$

These can be solved backward from V_m^T (taking $V_m^{T+1} \equiv 0$ and $c^{T+1,T+1} \equiv 1$ as boundary conditions). Then $f^{a,b}(x,y)$ can be recovered from Equation (3.3.3) using $s^{a,b}(x,y)$ along with the linear decomposition result:

$$\begin{aligned}f^{a,b}(x,y) &= s^{a,b}(x,y) - \rho d^{a+1,b+1}(x,y)s^{a+1,b+1}(x,y) \\ &- \delta d^{a,b}(x,y)c^{a+1,b+1}(x,y)\mu(\alpha^{a,b}(x,y)) - \rho(V_m^{a+1}(x) + V_f^{b+1}(y)) \\ &+ (r+\rho+\psi_m^a(x))V_m^a(x) + (r+\rho+\psi_f^b(y))V_f^b(y). \quad (4.4.2)\end{aligned}$$

Finally, the division of the marital surplus into the transfer terms

$$f^{a,b}(x,y) + z = t_m^{a,b}(x \mid y,z) + t_f^{a,b}(y \mid x,z)$$

can be recovered by using Equation (3.3.2).

$$\begin{aligned}t_m^{a,b}(x \mid y,z) &= (r+\rho+\psi_m^a(x))V_m^a(x) - \rho V_m^{a+1}(x) \\ &+ (1-\beta)\Big[(r+\rho+\delta+\psi_m^a(x)+\psi_f^b(y))\mathcal{S}^{a,b}(x,y,z) - \rho\mathcal{S}^{a+1,b+1}(x,y,z) \\ &- \delta\int \max\{\mathcal{S}^{a,b}(x,y,z),0\}dG(z)\Big]. \quad (4.4.3)\end{aligned}$$

Chapter 5

Results

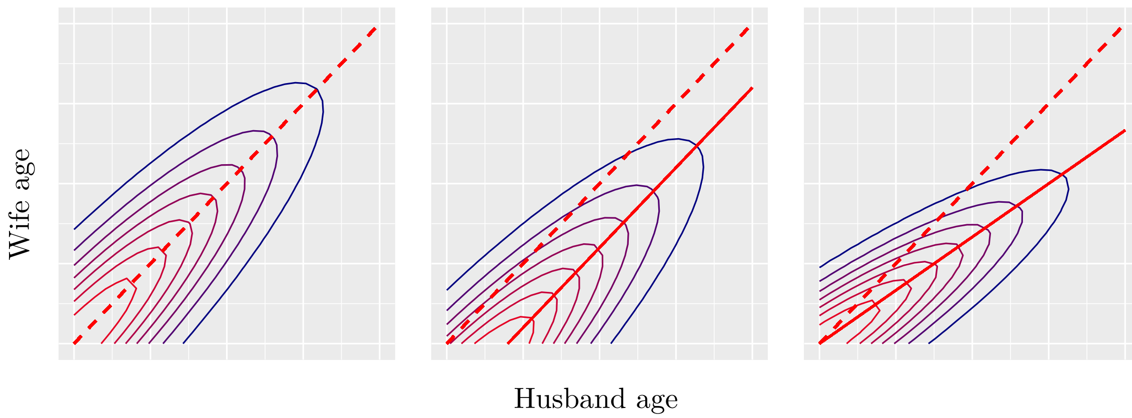
I first discuss the estimation results for the age-only model specification, and then proceed to the model with education and race types. The analysis is restricted to the top 20 largest MSAs, for which the arrival rates ξ and δ are estimated globally.

5.1 Age-only specification

The moment-based estimation procedure yields stable estimates of the arrival rate parameters. The estimate for the meeting arrival rate parameter is $\hat{\xi} = 2.59$ (with a standard error of 0.46). This is the arrival rate of meetings that singles experience when the overall sex ratio is balanced ($U_m/U_f = 1$). In other words, singles get dates on average every $2.59^{-1} \approx 0.39$ years (or 4.6 months).

The estimate for the arrival rate of love shocks is $\hat{\delta} = 0.0196$ (with a standard error of 0.0002). This means that love shocks get renewed on average every $0.0196^{-1} \approx 51.1$ years. This may seem too infrequent, but recall that arrivals are exponentially distributed, so that a large share of arrivals have much shorter durations. For example, the probability of an arrival within 7 years is 13%, which rises to 21% by 12 years. The median duration is 35 years. The model has to balance fitting both the marriages that dissolve early as well as those that persist for a lifetime. A higher arrival rate δ improves the fit for the former at the cost of worsening the fit for the latter.

The primary object of interest is the non-parametric estimate of the marital production function, $f^{a,b}$. Consider the possible shapes of $f^{a,b}$ and how they correspond to differential aging. Figure 5.1 shows three examples of what $f^{a,b}$ might look like.



Notes: The dotted line denotes $a = b$ for reference. The solid line denotes the transformation $a = b'$.

Figure 5.1: Examples of production function shapes

In these examples, marital output is both higher for couples of similar ages and for younger couples. The left panel shows the case in which $f^{a,b}$ is symmetric about the $a = b$ line of equal ages. This would mean that age affects the marital output identically for men and women. The middle panel shows the case in which $f^{a,b}$ is symmetric about a line $a = b + k$. This would mean that women are k years more mature than men — by redefining female age as the translation $b' = b + k$, we recover the case of symmetry about $a = b'$. This is a special case of differential aging: men and women mature at the same rates, but women have a k-year head start. A more interesting case of differential aging is when men and women mature at different rates. The right panel shows the case in which women mature $q > 1$ times faster than men. In this case, female age could be redefined as the rescaling $b' = qb$ to recover the case of symmetry about $a = b'$. Put another way, in order for women to mature more quickly than men, it would have to be the case that $f^{a,b}$ declines more rapidly in the direction of increasing b as compared to a. This is visible in the contours of the right panel: starting from a point on the solid line, moving upwards (increasing wife age) results in a steeper descent than moving rightwards (increasing husband age). In general, there will not be a linear transformation of age that makes the production function symmetric, but these cases provide some intuition for interpreting the shape of the production function.

Thus, a simple way to determine if men and women mature at different rates is to compare the rates at which marital output changes in a and b, i.e., the partial derivatives $\partial f^{a,b}/\partial a$ and $\partial f^{a,b}/\partial b$. This can be interpreted more intuitively as a hypothetical controlled experiment in which the age of one spouse is held fixed while the other spouse's age is manipulated in order to measure the change in the gains from marriage.

Figure 5.2 shows the global estimate for the marital flow output, $f^{a,b}$, obtained by averaging the estimates from each local marriage market (see Figure F.6 in Appendix F for the estimates in each of the 4 largest MSAs). The shape of the surface has two interesting features. First, marital output is very low for young people but rises rapidly in the mid-twenties. Recall that static matching models underestimate the gains from marriage for young people, as they do not account for the value of search over the lifecycle. Thus, even taking into account the value of search with a dynamic lifecycle model, the gains from marriage are generally very low for young people. Second, marital output is highest along the diagonal of similar ages, and falls off sharply with large age gaps. This implies that age similarity is a major driver of the gains from marriage, and not just an outcome of equilibrium sorting. Compare the marital production function to the distribution of marriages (the equilibrium matching outcome) in Figure 4.1: the latter is much more tightly concentrated around the diagonal. This is not surprising, however, as equilibrium sorting often exaggerates the shape of the underlying production function.

As for differential aging, the marital output is highest when the husband is slightly older than the wife (although this reverses at later ages). But there is no line of symmetry: the contours bulge out more below the $a = b$ line, indicating that marital output does not drop as much when the husband is older than the wife. This reflects differential aging in the sense that women mature at a faster rate: female aging reduces marital output more than does

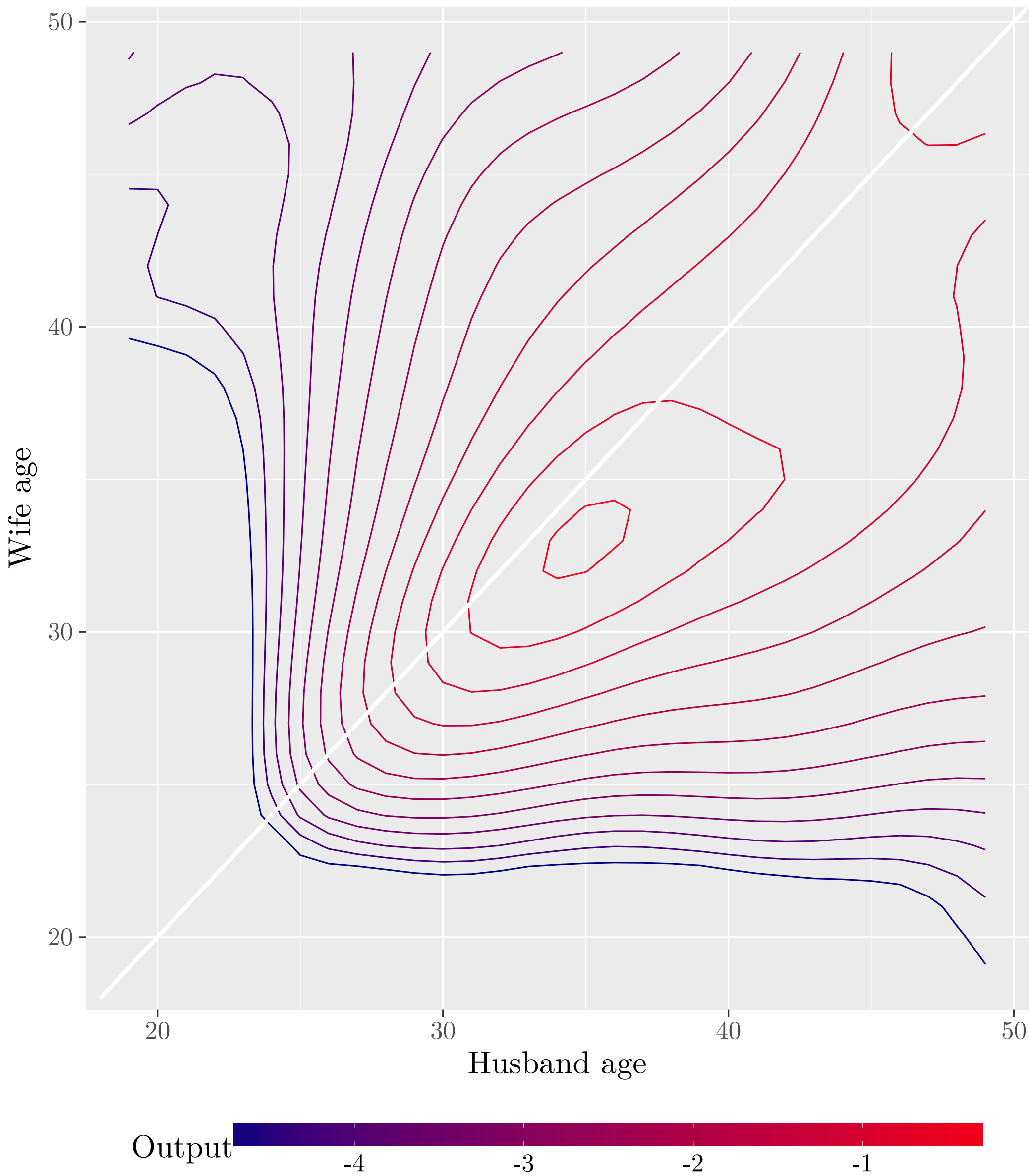
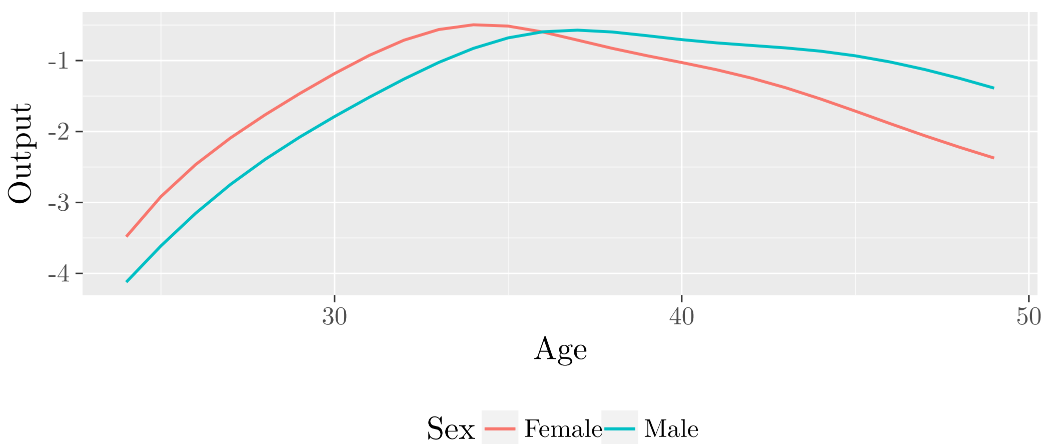


Figure 5.2: Marital flow output

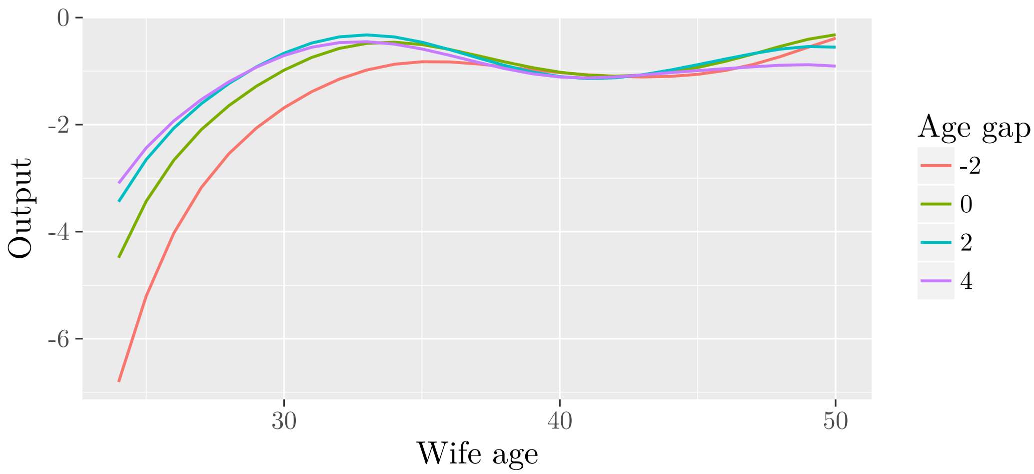
male aging. For example, a 36 year-old woman would be best matched with a 38 year-old man, but starting from this point, there is a greater reduction in output from increasing the wife's age than the husband's. The magnitudes of these productivity changes are more easily seen in Figure 5.3, which shows slices of the production function where one spouse's age is held fixed. The curves show how the marital output for a given person varies with the age of their spouse. In this case, when the spouse's age is increased past the point of maximum productivity, the rate of dropoff is roughly twice as rapid in the wife's age as compared to the husband's age.



Notes: The line for females is the output when married to a 38 year-old man. The line for males is the output when married to a 36 year-old woman.

Figure 5.3: Marital flow output slices

Though this is evidence of differential aging, it is not clear that marital output is strongly affected by female fecundity in particular. If it were, it would be visible as a steeper dropoff in output as the wife age increases from mid-thirties to early forties, which is the period over which female fecundity sharply declines. However, the decline in marital output is not noticeably steeper in that region, which suggests that differential aging is not driven primarily by differential fecundity. On the contrary, the effect of fecundity may be obscured by the presence of children. As children increase the gains from marriage, this could mask the impact of fecundity in the estimate of the marital production function. This possibility is discussed further in Section 5.3.2 below.



Notes: A two-year age gap means that the husband is two years older than the wife, i.e., $a = b + 2$.

Figure 5.4: Marital flow output for couples with a two-year age gap

Another way to look at marital output is to see how it changes over the course of a couple's lifecycle. Figure 5.4 shows the path of marital output for couples of various age differences. In general, marital output increases very steeply through the twenties, peaks in the early thirties, and remains high into old age. There is also a notable dip during the forties, which is likely an artifact of the differences in marital output across education levels, as will be seen with the other specification in Section 5.2.

5.1.1 Discussion

The results show that marital output declines more sharply in female age than male age. For a typical couple in their late 30s, marital output declines roughly twice as rapidly in female age. This provides unambiguous evidence for the differential aging hypothesis — specifically, that women mature faster than men. As the model accounts for equilibrium sorting, it effectively controls for sex ratio effects in order to estimate the marital output (which is a model primitive and independent of equilibrium effects).

This model specification only accounts for assortative matching by age (and MSA), but there are other important dimensions along which people sort, such as education and race. Sorting on unobservables would introduce bias into the estimation, as the equilibrium model would be misspecified. For example, if college graduates strongly prefer to marry other college graduates, then the sex ratio within educational classes becomes important in the equilibrium, even though the overall sex ratio may be balanced. If the supply of college-educated men is exceeded by the demand from college-educated women, then some of these women will have to either marry outside of their educational class or remain single. As a robustness check, I address this limitation in Section 5.2 below by including education and race in the agent type vectors.

5.1.2 Other non-parametric estimates

Figure 5.5: Value of singlehood

Figure 5.5 shows the average value functions of singlehood for men and women in the four largest marriage markets. The option value of search is high for the young, and declines steadily with age.[45] In each market and at almost every age (except for the youngest

[45]There is likely some boundary bias at the youngest ages, resulting in an underestimate of the value of

people), men have a higher value of being single and searching.[46] This indicates that men have more favorable marriage opportunities in these marriage markets. This is no surprise, as the sex ratios in these marriage markets become ever more favorable to men as they age (see Figure 2.3 above). With their marriage prospects improving, men have more incentive to wait for a good match. For women, on the other hand, not only does the sex ratio work against them, but differential aging further penalizes them for delaying marriage. Thus, women have strong disincentives to being single and searching.

Figure 5.6 shows estimates of $\alpha^{a,b}$, the probability of marriage conditional upon meeting, in the four largest marriage markets. Although marriage is strongly positively assortative on age, there is also a very prominent tendency toward marriages in which the woman is younger than the man. First of all, match probabilities are highest where women are 1–2 years younger than men. But, as with the marital output, the contours bulge out more below the $a = b$ line and so the match probability declines more gradually in the man's age than in the woman's age. $\alpha^{a,b}$ embodies equilibrium matching strategies, so it takes into account both the marital output $f^{a,b}$ as well as the relative population supplies of men and women in the local marriage market. The asymmetry is more pronounced here because the effects of sex ratio imbalance are combined with the effects of differential aging.

5.1.3 Model fit

The model is fit to data on marriage and divorce flows: the estimator minimizes the deviations between the flows in the data and those implied by the model. The estimation chooses δ and ξ so as to best fit both the marriage and divorce flow moments. The mechanics of the fitting these two parameters can be interpreted from Equation (4.3.2). A larger δ generates

search. This is because agents in the model who are near the age boundary cannot date people younger than the age cutoff, whereas in reality they can and do. This makes it appear as though the value of search is lower near the age boundary. The true value functions are more plausibly monotonically decreasing.

[46] Some cities have lower values of search overall. This is an artifact of the estimation procedure — the non-parametric estimates, including the value functions of singlehood and the marital production functions, are calculated separately for each MSA. As such, cities may have higher or lower estimated marital productivities, which are reflected in the value function estimates. The marital production function is a model primitive and should be the same everywhere if the model is correctly specified. To arrive at a global estimate of the marital production function, I average together the local estimates.

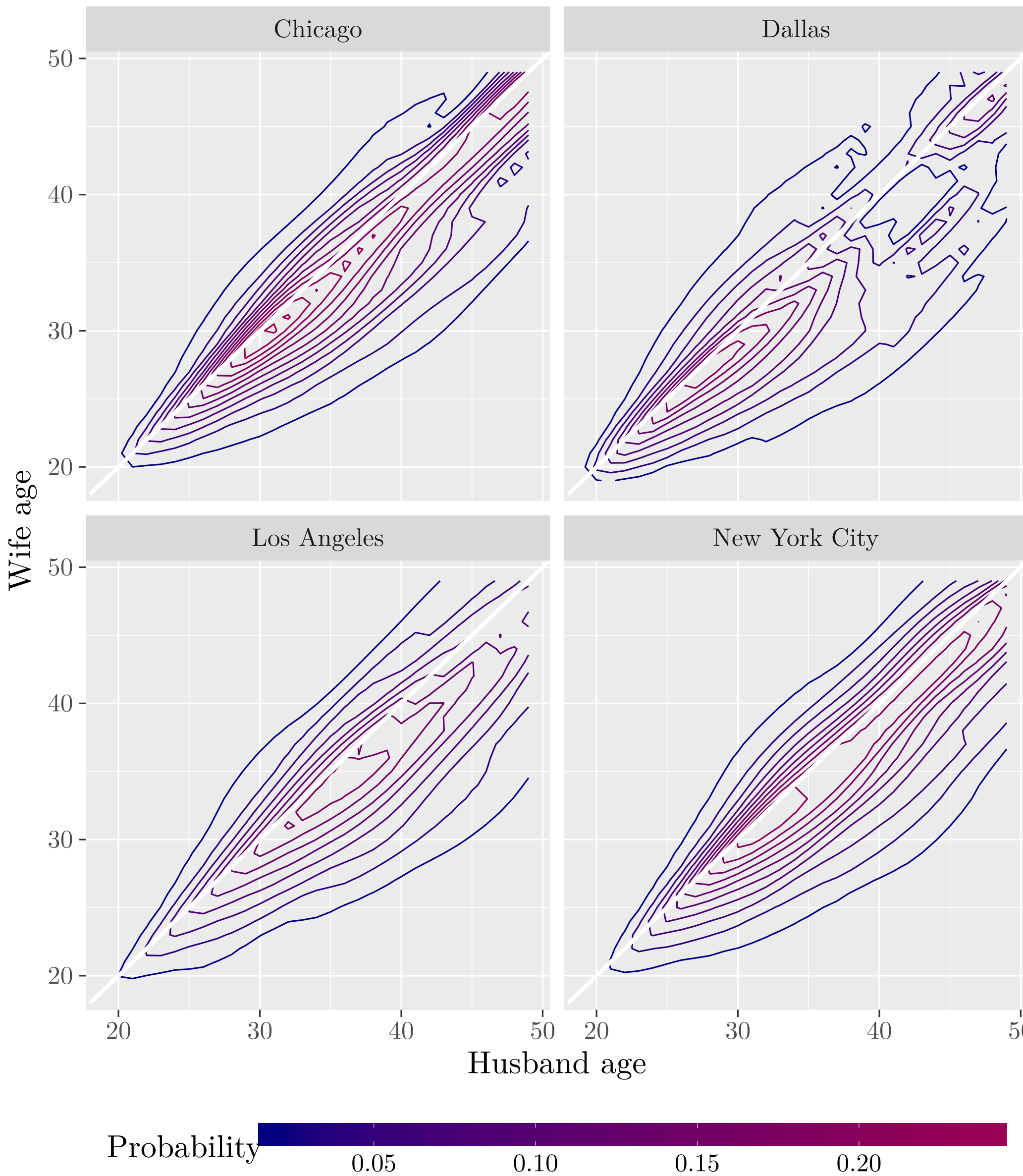


Figure 5.6: Marriage probability conditional on meeting

more divorces, but also has the effect of increasing the flow of marriages, as α increases in δ! This is because increasing δ increases both the numerator and the denominator by the same amount, but since α is a probability, the numerator cannot be larger, and hence the ratio must increase. Intuitively, marriages are less durable, and so the returns to searching for a better match are smaller, making present matches more enticing. A smaller ξ generates fewer marriages, but also increases α, which has the effect of decreasing divorce flows. Intuitively, meetings are harder to come by, and so the returns to searching are smaller, making divorce less enticing.

Notes: Solid lines indicate smoothed counts from the data. Dashed lines are fitted values from the model. Curiously, there are consistently more female divorces in the sample, which seems to violate the accounting identity that the total number of divorces should be equal for men and women. Aside from the possibility of sample bias or misreporting, this could be a result of men being more likely to relocate out of the MSA following a divorce, or of women being more likely to move to a large MSA following a divorce.

Figure 5.7: Model fit for divorce flows

Figure 5.7 shows the smoothed divorce flows from the data along with those implied by the fitted model for the four largest marriage markets, the metropolitan areas of New York

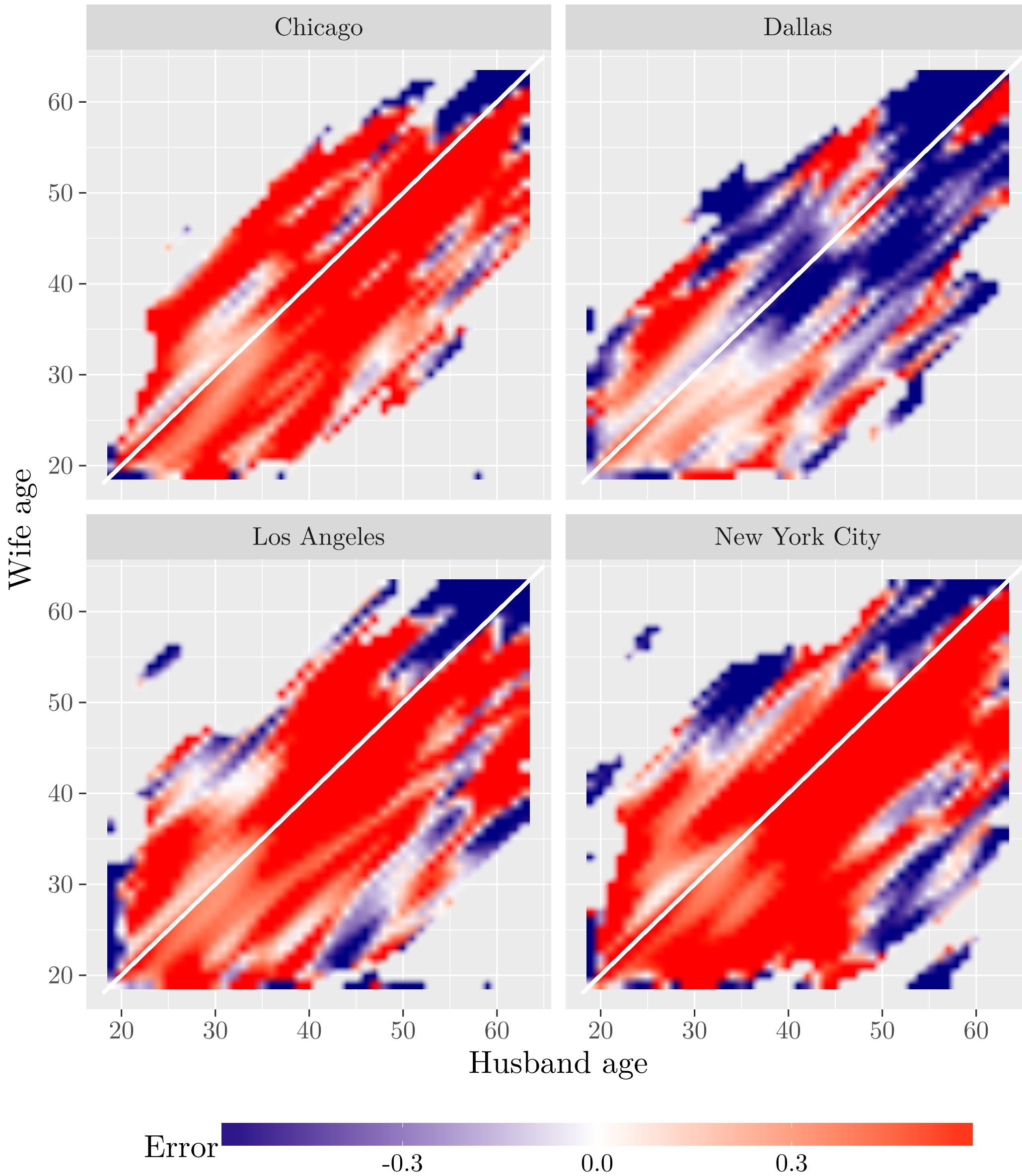
City, Los Angeles, Chicago, and Dallas. The model moments exhibit a particular pattern: they underpredict divorce for younger people and overpredict it after age 50. As will be discussed in Section 5.3.1 below, this difficulty in fitting the overall pattern of divorce flows is a product of the assumption that love shocks are i.i.d. as opposed to serially correlated. In short, the love shock process cannot generate enough long-duration marriages to match the low divorce flows after age 50, and so the estimation splits the difference by setting $\hat{\delta}$ to an intermediate value, where it balances the error from underpredicting divorce for the young with the error from overpredicting divorce for the old.

Figure 5.8 shows the difference between the marriage flows implied by the fitted model and those observed in the data, for the four largest MSAs. In these cities, the model fits relatively well in the regions with the largest marriage flows, i.e., below age 35. Outside of this region, the model fit is poorer, understating the marriage flows in Dallas and overstating them in the other three largest MSAs (though keep in mind that the model parameters are fit on the 20 largest MSAs).

5.2 Age, education, and race specification

Adding education and race as multidimensional types allows the model to more closely capture the matching patterns on the marriage market. With two education and two race categories, there are 16 possible kinds of matches for each age pair. This results in a data sparsity problem, especially for types where there are very few observed marriages. To alleviate this, I increase the smoothing parameter for the non-parametric regression of the marriage stocks, h_1, from 16 to 24.

Now the estimate for the meeting arrival rate parameter is $\hat{\xi} = 1.56$ (with a standard error of 0.12). In other words, singles get dates on average every $1.56^{-1} \approx 0.64$ years (or 7.7 months). This is less frequent than in the first specification, which indicates that the model can more easily generate the observed marriage flows by accounting for assortative mating on race and education. The estimate for the arrival rate of love shocks is $\hat{\delta} = 0.0215$ (with a standard error of 0.0003). This means that love shocks get renewed on average



Notes: Error is calculated as the proportional error between the model and the smoothed data counts, with the smoothed data counts taken as the basis. Age pairs with insufficient counts in the smoothed data are excluded.

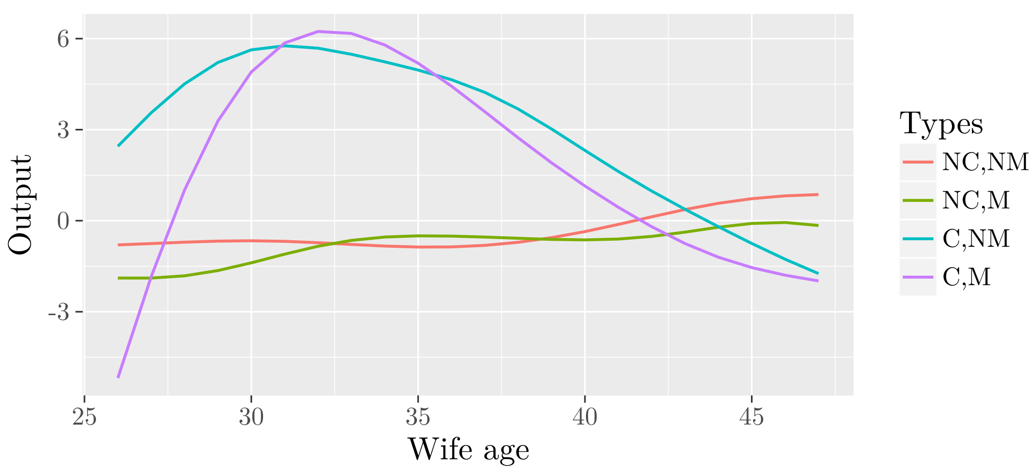
Figure 5.8: Model fit for marriage flows

every $0.0215^{-1} \approx 46.4$ years. Table 5.1 shows the estimates under both specifications for comparison, with standard errors in parentheses.

Specification	$\hat{\xi}$	$\hat{\delta}$
Age only	2.59 (0.46)	0.0196 (0.0002)
Age, education, and race	1.56 (0.12)	0.0215 (0.0003)

Table 5.1: Arrival rate parameter estimates

The results concerning differential aging do not qualitatively change; several plots are provided in Appendix F. Figures F.7 to F.10 show the global estimates for the marital flow output for all four homogamous type pairs. Although the gains from marriage differ by type pair, the pattern of marital output dropping less in the husband's age remains very prominent. The shapes of the marital production functions for homogamous type pairs reveal important differences. Figure 5.9 shows the $a = b + 2$ slices of the marital output for each homogamous type pair, i.e., how the flow output changes over the lifecycle for couples in which the wife is two years younger than the husband. Figures F.11 to F.14 show the value functions of singlehood for each type.



Notes: A two-year age gap means that the husband is two years older than the wife, i.e., $a = b+2$. Education types are denoted by C/NC (college/non-college). Race types are denoted by M/NM (minority/non-minority).

Figure 5.9: Marital flow output for couples with a two-year age gap

5.2.1 Education

There is a striking difference in the gains from marriage between education levels, mirroring the differences in marriage rates shown in Figure 2.6. For college-educated couples, marital output soars in the early thirties before dropping back to the non-college baseline by the forties. In other words, the estimated model explains the much higher marriage rates of college-educated people as the result of them enjoying much larger gains from marriage. This sharp divide along educational class lines reflects a pattern discussed by Lundberg and Pollak (2013): college graduates generally use marriage as an arrangement for making large investments in their children, such as saving for their college tuition, whereas less-educated people generally do not. As such, the college-educated are much more likely to marry before having children, and less likely to separate than those without a college degree. The spike in marital output coincides with the period of childbearing and raising young children, which is consistent with the hypothesis that, for college graduates (who follow the high-investment strategy), there is a lot of value in raising young children within marriage.[47] This is perhaps because fathers want to establish their parental rights (and a high degree of certainty of their paternity) before committing to making large parental investments.[48] Moreover, Chiappori, Salanié, and Weiss (2017) finds that the gains from marriage have sharply increased for college-educated women since the 1970s, and that college-educated parents are now spending much more time with their children.

Unsurprisingly, the value of search is much higher for college graduates, especially before the age of 35, when their marital output is very high. The value functions also reflect

[47] A potentially confounding factor is the labor market, which is omitted from the model. It is plausible that college graduates face different career incentives that lead them to delay marriage until their late twenties or thirties. For example, it may be very advantageous for a young college graduate to be highly mobile and able to easily relocate to pursue lucrative career opportunities. The result of this is to concentrate marriage activity into this age window. In my model, this would bias the estimates of marital output upwards in this age window, as the model would only be able to generate those large inflows into marriage via higher marital output at those ages. See Flabbi and Flinn (2015) for a model of simultaneous labor and marriage search.

[48] Edlund (2013) posits a theory of marriage (as opposed to mere cohabitation) as an exchange in which the wife transfers (some or all) custodial rights over her children to her husband in return for his contribution of financial resources. As there is no market for children, men have no choice but to marry in order to acquire parental rights.

the effects of the sex ratio imbalances stemming from higher female college attainment, as discussed in Chapter 2: among college graduates, men have a much higher value of search than women. For those without a college degree, at least for non-minorities, men and women have roughly equal values of search — for minorities, sex ratio imbalances play a large role as discussed below.

Note however that these estimates are based on the assumption of stationarity, when in fact there have been large increases in educational attainment over the past several decades, especially for women. This is visible in how the population supplies in Figures F.2 and F.3 vary by cohort: the populations of college graduates have been growing with each new cohort. This biases the estimation when education is included as a static type, and so the results should be taken with a grain of salt. Consider that $\alpha^{a,b}$ is estimated as a first difference of the form $m^{a,b} - m^{a-1,b-1}$, from Equation (4.3.2). In a world in which educational attainment is rising, this biases the estimates for $\alpha^{a,b}$ downwards for the college-educated (and upwards for the less-educated). Marital output $f^{a,b}$ is generally increasing in $\alpha^{a,b}$, and so this bias propagates through to the estimated gains from marriage. Beyond changes in educational attainment, any non-stationarity in the demographics of a city will bias the estimation, which I discuss further in Section 5.3.4.

5.2.2 Race

Racial differences in marital output are more muted by comparison. For those without a college degree, the marital output for minorities drops significantly lower after the age of 45, and slightly lower before the age of 30, relative to non-minorities. Figure 2.7 shows that marriage rates for blacks are much lower than for whites or Asians. Seitz (2009) summarizes the literature which finds that, compared to whites, blacks are less likely to marry, and more likely to divorce. The fact that marital output is not much lower for minorities between the ages of 30 and 45 suggests that equilibrium sorting is the source of these outcomes. Indeed, the population supplies for minorities in Figures F.4 and F.5 show that there are quite extreme sex ratio imbalances. After the age of 45, the lower marital output indicates that sex ratio imbalance alone cannot account for the differences in marriage and divorce. To this point, past the age of 50, the sex ratio imbalance is at least as bad for non-minorities. For college graduates, the overall shape of marital output over the lifecycle is very similar by race. Due to data sparsity for college-educated minority couples, it is not evident that there are any meaningful racial differences in the marital output.

The value functions of search reflect the relative shortage of minority men (in the large MSAs studied herein), as discussed in Chapter 2: among people without a college degree, non-minority men and women have roughly equal values of search whereas minority men have a much higher value of search than minority women. It is also noteworthy that the value of search is much lower for minorities. Given that their marital output is not much different from that of non-minorities, this is likely an artifact of the meeting technology. With meetings being drawn at random from the population, minorities do not meet each other very often, which reduces their value of search insofar as they prefer to marry one another.

5.3 Limitations

The model is a greatly simplified representation of the real world process of meeting and marrying, and it only aims to capture the broad patterns of marriage. There are a number of limitations which prevent the model from more accurately fitting the data.

5.3.1 Divorce process

A central limitation of the model is that the renewals of love shocks are i.i.d., which is a very restrictive assumption. As such, a large love shock provides no basis for a durable marriage, as love shocks are independent of one another. A much more realistic assumption is that love shocks are serially correlated. This neatly captures the intuition that there are many idiosyncratic factors that contribute to marriage quality, most of which are typically unobservable to the researcher (e.g., personality, habits, values, etc.), but which are enduring factors that can support durable marriages. As there will always be some unobserved heterogeneity which affects marriage decisions, modeling the persistence of such factors is an important objective for any model of divorce.

Because of this limitation, the model has trouble matching the distribution of marriage durations in the data. It cannot simultaneously generate enough lifelong marriages and enough divorces, as the arrival rate parameter δ is the only means of adjusting the divorce rate. This is visible in how the model fits the marriage and divorce flows. For example, the model cannot simultaneously generate both small marriage flows and small divorce flows, as happens after age 50. Small flows into marriage imply a low α, whereas small flows out of marriage imply a high α. As a result of this tension, the model underestimates the divorce rate for the young and overestimates it for the old.

In other words, the model generates a distribution of marriage durations (conditional on observed types) with a thinner tail than is observed in the data. With persistent love shocks, the model could more easily generate long-duration marriages, even with a larger value of δ (i.e., more frequent arrivals of love shocks). The extra persistence parameter introduced by a Markov process would be identified from data on divorce flows, as it would directly affect the shape of the distribution of marriage durations. However, adding persistence to the love shock process requires keeping track of the current draw as a state variable, and so the current method of solving the model no longer works. Rıos-Rull, Seitz, and Tanaka (2016) make use of a simple two-state Markov process although they simulate their model for estimation. Incorporating persistence into the love shock process is a significant but crucial challenge for future research.

5.3.2 Children and parenting

Another significant limitation is that the model abstracts away from impact of raising children on the gains from marriage. Female fecundity is surely a dominant component of marital output for childless couples. For couples who already have children, female fecundity is much less important; for couples who do not desire any additional children, it is completely irrelevant. In this case, the gains from marriage may largely derive from raising the children together within the family unit. It could be the case that marital output falls very sharply for unmarried childless women at the end of their fecund years, but does not drop at all for married mothers of the same age.

My model does not distinguish between these two cases and so I only recover an overall estimate of the marital output, regardless of whether or not children are present. The presence of children ought to be represented in the model by a state variable so that it is factored into the marital production function. If raising children together substantially increases the marital output, then couples with children will be less likely to separate. Thus, because the presence of children affects divorce flows, the marital output is separately identified for couples with and without children.[49] By estimating the gains from marriage separately for couples with and without children, one could identify the impact of differential fecundity and the extent to which it explains my differential aging result.

This also raises interesting questions around step-parenthood. Divorced mothers who cannot have (or do not want) additional children may not be very desirable on the marriage market, which could explain the lower remarriage rates for women. In fact, they may be less desirable than comparable childless women because step-parenthood could be a deterrent to men, who would prefer not to expend their resources raising biologically unrelated offspring.

5.3.3 Cohabitation

A potential confounder of marriage studies is the ubiquity of cohabitation as a substitute for marriage. Much of the gains from marriage can be captured through long-term cohabitation, such as sharing a residence or having children. As Lundberg and Pollak (2013) document, over the last half-century, cohabitation has gone from being very rare to very common. However, there is an educational divide in cohabitation patterns. College graduates mostly cohabit as an intermediate step on the road to marriage, but less-educated people are to a large extent forgoing marriage entirely. As my estimation treats cohabiting

[49]With ACS data, the challenge lies in attribution of divorces to marriages with and without children. It is only possible to observe children within a household, and so only the parent with custody would be identifiable as having exited a marriage with children. The other parent would indistinguishable from a childless divorcee as it is not observable whether or not they are a parent.

couples as singles, the estimate of marital output is biased downward for couples without a college-education. Brien, Lillard, and Stern (2006) estimate a search model of the marriage market that allows for an intermediate cohabitation stage as a precursor to marriage. This allows the pair to learn about one another before committing to marriage. Accounting for cohabitation allows the estimation to exploit additional information on match formation prior to or in lieu of marriage.

5.3.4 Non-stationarity

Finally, stationarity is a restrictive assumption, albeit a central one for solving the model. The model cannot account for major demographic and social shifts such as declining fertility or the large increase in female college attainment and labor force participation. Moreover, with estimation being done at the MSA level, any generational demographic shifts within cities will bias the estimates. However, given the difficulty of relaxing this assumption, it is not clear if much progress can be made in this direction. To the best of my knowledge, every dynamic equilibrium matching model in the literature relies on the steady-state assumption.

Chapter 6

Conclusion

In this paper, I investigate the sex differences in marriage patterns over the lifecycle. In short, women are more likely to marry young, whereas men tend to marry later and are much more likely to remarry. On one hand, these patterns may stem from fundamental differences between men and women. If aging affects men and women differently, or equivalently, if men and women have differing marital preferences with respect to spousal age, then observed marriage patterns may simply reflect these fundamental differences. I call this mechanism "differential aging". For example, a number of papers have argued that the shorter female fecundity window (i.e., differential fecundity) can explain a wide range of the sex differences in marriage outcomes. On the other hand, it is also possible that equilibrium sorting could be a contributing, or even dominant factor, in producing these sex differences. Differing population supplies of men and women can produce sex ratio imbalances which have quite extreme impacts on equilibrium matching outcomes.

To determine the extent to which differential aging accounts for these differences, I develop an equilibrium model of lifecycle search and matching in order to estimate the gains from marriage as a function of husband and wife ages. As the model accounts for equilibrium sorting in the marriage market, it can be estimated to recover the underlying marital output, which represents the gains from marriage. I find that the marital output for a typical couple in their late 30s falls twice as rapidly in the wife's age as in the husband's age. This indicates that differential aging plays a fundamental role, generating asymmetric gains to marriage in terms of aging for each sex. Thus, the lower female marriage rates at later ages are to a large extent due to differential aging, i.e., lower gains from marriage. However, it does not appear that differential fecundity in particular is a major contributor to differential aging, although this is difficult to detect as the model does not account for children.

I also find very large differences in marital output with respect to education. College graduates experience extremely large gains from marriage compared to the non-college population, at least before age 40. This explains the much higher prevalence of marriage among the college-educated. By comparison, differences in marital output between races are relatively small, indicating that equilibrium sorting is the dominant factor in explaining the racial differences in marriage rates. Minorities without college degrees experience substantially lower gains from marriage at later ages, suggesting that unbalanced sex ratios do not fully explain their lower marriage rates. However, sex ratios among minorities are still quite unbalanced, and so policies targeted at helping disadvantaged men will indirectly help their female counterparts by alleviating the shortage of marriageable men. On the other hand, college graduates experience similar gains from marriage regardless of race.

I show how to identify the structural model from cross-sectional data, allowing for the use of datasets large enough to estimate richer models. For instance, I estimate the model at the MSA level — a much more realistic geographic delineation of a marriage market — which allows me to capture the effects of local sex ratio imbalances on equilibrium sorting. I also perform an additional estimation of a model specification that includes binary education and race types (alongside age). To deal with problems of data sparsity, I employ a non-parametric regression to smooth the data. The ability to identify marriage market models from large cross-sectional surveys is a boon to future research in the area, as sample size limitations tend to be the limiting factor in estimating richer marriage market models. A promising avenue for future exploration is to incorporate major lifecycle events into the model, such as fertility and children, education and career choice, and unemployment.

Many of these factors have potentially large effects on the marital output and hence on the distribution of observed marriages. Thus, there is much to learn by explicitly modeling them.

Finally, there are two important challenges for future research on marriage markets within the search framework. The first is to generalize the divorce process to allow for serially-correlated love shocks. This is essential for accurately modeling unobserved heterogeneity in matching outcomes. By allowing for serial correlation, the model can much more accurately fit the observed distribution of marriage durations. The second challenge is to incorporate fertility into the model by explicitly modeling the arrival of children. Estimating such a model would allow for the separate identification of the gains from marriage for couples with and without children. Doing so is essential for isolating the impact of differential fecundity, which would primarily be observable in the marriage and divorce patterns of childless women.

Appendices

A Estimation details

All code used in the estimation is available in a Git repository online at `https://github.com/tobanw/search-match-age-estimation/`. The `dissertation` branch preserves a snapshot of the code used in producing the published dissertation.

A.1 Data preparation

To compute population stocks and flows in each MSA, I pool the 2008-2016 ACS 1% surveys (9 years of data) from Ruggles et al. (2017). The dataset provides sampling weights for both individuals (`PERWT`) and households (`HHWT`) so that nationally representative statistics can be computed (as some groups are oversampled). Anyone living in group quarters (i.e., prisons and mental institutions), denoted by the variable `GQ`, is excluded. MSA codes are taken from `MET2013`. College completion is defined as having completed at least 4 years of college, i.e., `EDUC` ≥ 10. This is not perfectly accurate, but there is no variable in the ACS data indicating whether someone has earned a college degree. Race is classified into two groups: minority and non-minority. Non-minority is defined as declaring a primary race of white or Asian, determined from `RACED`.[50] Marital status is determined from `MARST`. To count married couples, I tally over married women, keeping track of husband attributes. The population stocks and flows used in the model are computed in each MSA as follows:

- Population stocks for individuals, $\ell_m^a(x), \ell_f^b(y)$: counts weighted by `PERWT`
- Population stocks for singles, $u_m^a(x), u_f^b(y)$: counts weighted by `PERWT`
- Marriage stocks for couples, $m^{a,b}(x, y)$: counts weighted by `HHWT`
- Marriage flows for couples, $MF^{a,b}(x, y)$: marriage in the past year is determined from `MARRINYR`; counts weighted by `HHWT`

[50]The Asian category is broadly defined and includes Indians as well as Middle-Easterners. The exact categorization that I used for non-minorities is any `RACED` code among the following: 100-130, 400-699, 810-826, 860-892, 910-925, 943, and 963. Anyone else is categorized as a minority.

- Divorce flows for individuals, $DF_m^a(x), DF_f^b(y)$: divorce in the past year is determined from `DIVINYR`; counts weighted by `PERWT`

- Net migration outflows of married couples, $\varphi^{a,b}(x,y)m^{a,b}(x,y) - \gamma^{a,b}(x,y)$: migration in the past year is determined from `MIGRATE1D`, and the source MSA can be recovered using `MIGPUMA1` and `MIGPLAC1` (but this only works in survey year 2012 and onward); counts weighted by `HHWT`

These are then divided by the number of survey years used in counting to get true population and annual flow estimates. Subsequently, smoothing is applied as described in Section 4.2.

A.2 Optimization details

For the arrival rate parameter estimation, I used the COBYLA optimizer described in Powell (1998).[51] I also supplied non-negativity constraints to the optimizer. The GMM objective function is well-behaved with a unique global minimum, so the optimizer converges fairly quickly.

A.3 Bootstrap standard errors

To calculate standard errors for the arrival rate parameter estimates, I performed the estimation on 100 bootstrap samples and computed the empirical standard deviation of the resulting estimates. Bootstrap resampling was performed at the household level within each MSA and survey year in order to more closely replicate the ACS sampling scheme. The ACS is a nationally representative sample of 1% of the US population in each survey year; the sampling unit is the household and the sampling is stratified by region.

I used GNU Parallel[52] to batch process the smoothing and estimation steps in parallel on multiple processors.

[51] Specifically, I used the implementation of COBYLA provided by `NLOpt` 2.4.2, Johnson (2014).
[52] Tange (2018)

A.4 Computational details

When G is $\mathcal{N}(0, \sigma^2)$, $G(z) = \Phi(z/\sigma)$ and

$$\mathbb{E}\left[z \mid z > -s(x,y)\right] = \frac{\sigma\phi(-s(x,y)/\sigma)}{\alpha(x,y)}.$$

So in this case,

$$\mu(a) = \sigma\left(\phi(\Phi^{-1}(1-a)) - a\Phi^{-1}(1-a)\right).$$

A.5 Arrival rate parameter regression

Goussé, Jacquemet, and Robin (2017b) employ the following estimation strategy for λ and δ. Dividing the marriage and divorce flows in Equation (4.3.1) by population stocks yields marriage and divorce rates

$$\begin{aligned} MR^{a,b}(x,y) &:= \frac{MF^{a,b}(x,y)}{u_m^a(x)u_f^b(y)} = \lambda\alpha^{a,b}(x,y) \\ DR^{a,b}(x,y) &:= \frac{DF^{a,b}(x,y)}{m^{a,b}(x,y)} = \delta(1-\alpha^{a,b}(x,y)). \end{aligned}$$

Then α can be eliminated by summing these two equations, yielding a reduced form equation that can be estimated by OLS:

$$1 = \frac{1}{\lambda}MR^{a,b}(x,y) + \frac{1}{\delta}DR^{a,b}(x,y). \tag{A.1}$$

However, this method requires that $DF^{a,b}(x,y)$ is observable — in the ACS data, only $DF_m^a(x)$ and $DF_f^b(y)$ are observable because there is no way to find the attributes of the ex-spouse. As only aggregate divorce flows are required for estimation, one possible workaround is to indirectly infer divorce flows from the attrition in marriage stocks. Obtaining accurate estimates of pairwise divorce flows $DF^{a,b}(x,y)$ requires very accurate measurements of marriage stocks $m^{a,b}(x,y)$ in each year and in each MSA. In my case, the data became too sparse at this level of granularity to provide usable divorce flows, but I outline the approach below for completeness.

Inferring divorce flows

Inferring $DF^{a,b}(x, y)$ from multiyear data on marriage attrition within cohorts is hypothetically possible if marriage stocks $m^{a,b}(x, y)$ are measured with sufficient accuracy from year to year. Within each cohort, I compared the stock of marriages in a year with the stock of marriages in the following year, excluding new marriages. This gives an estimate of the attrition from marriage. After netting out expected deaths, this gives an estimate of divorce flows.

Unfortunately, performing the regression using couples' marriage and divorce flows gives poor estimates of the arrival rate parameters. Figure A.1 shows the regression line plotted against the flows data. Data points are weighted by the corresponding stock of marriages. Note that the estimated divorce flows are very noisy and many data points are negative.

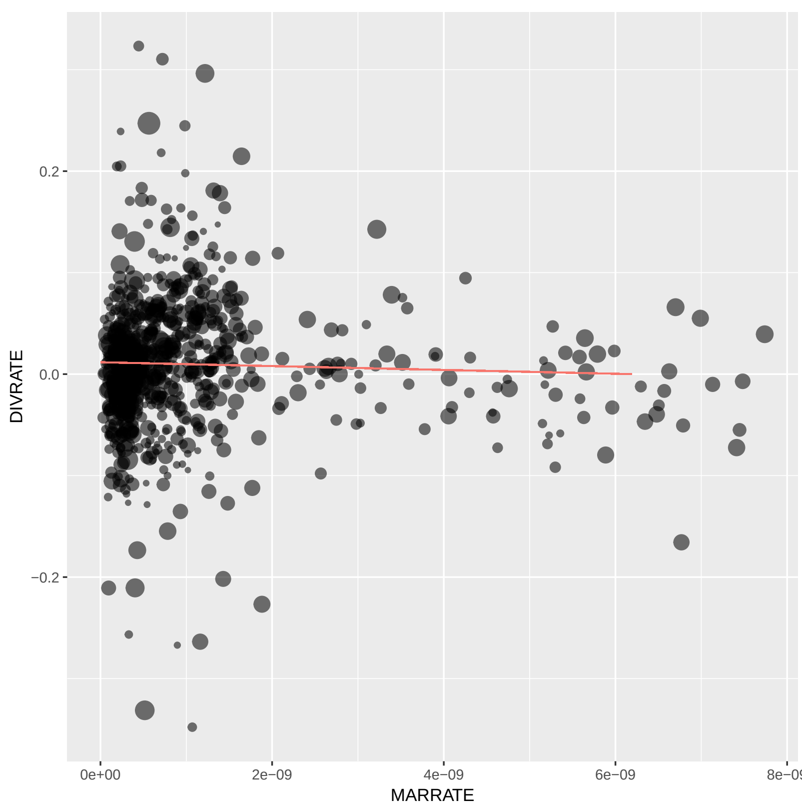


Figure A.1: Arrival rate parameter regression

B Model simulation

Solving the model for simulation purposes can be done by means of a nested fixed-point algorithm. The inner fixed-point iteration finds a matching equilibrium, taking the current measures of singles and couples as given. The outer iteration then updates these population measures to the values implied by the matching equilibrium until a market equilibrium is reached (i.e., the matching equilibrium generates a steady state).

Given α generated by a matching equilibrium, the singles and couples measures can be updated by solving the system of equations given by the steady state condition of Equation (3.5.1). (Without aging, the singles measures could be computed directly by integrating over the opposite sex.)

Given singles measures, a matching equilibrium is computed by fixed-point iteration. Starting with an initial guess for α, compute the singlehood value functions using Equation (4.4.1). Then solve for s using Equation (4.4.2) along with the value functions and α. Finally, use Equation (3.4.2) to update α for the next iteration.

Thus, the algorithm to compute a market equilibrium is a nested fixed-point iteration:

1. Start with an initial guess for α.

2. Use α to update $u = (u_m, u_f)$ by solving the steady state flow condition of Equation (3.5.1).

3. Solve the nested matching equilibrium given u, and α:

   (a) Solve Equation (4.4.1) for the implied value functions.

   (b) Solve Equation (4.4.2) for the implied surplus.

   (c) Use the surplus to compute the implied α from Equation (3.4.2).

4. Repeat with updated α until convergence.

The convergence criterion is that successive singles measures converge:

$$\max\{|u^{(k+1)} - u^{(k)}|\} \leq \varepsilon.$$

In practice, the fixed-point operator is not globally contracting and the algorithm will often get stuck in an oscillating pattern and fail to converge. This can be overcome through the use of a damping factor to shrink the α update:

$$\alpha^{(k+1)} = 0.8\alpha^{(k)} + 0.2\alpha^{(k')}.$$

C Derivation of value functions

I will solve for the value functions as limits of a discrete time model with a small time increment dt. In discrete time, the arrival rate parameters represent probabilities of an arrival per unit of time. For example, $\rho\, dt$ is the per-period probability of an aging shock arrival.

C.1 Value function for singles

For a single man of age a and type x,

$$\begin{aligned}(1+r\,dt)V_m^a(x) = (1-\psi_m^a(x)\,dt)\Bigg\{&(1-\rho\,dt)\times \\ &\Bigg[dt\sum_{b=1}^{T}\iint \lambda^{a,b}(x,y)\max\{W_m^{a,b}(x \mid y,z), V_m^a(x)\}u_f^b(y)dG(z)dy \\ &\quad + \left(1-dt\sum_{b=1}^{T}\int \lambda^{a,b}(x,y)dy\right)V_m^a(x)\Bigg] \\ &+\rho\,dt\Bigg[dt\sum_{b=1}^{T}\iint \lambda^{a+1,b}(x,y)\max\{W_m^{a+1,b}(x \mid y,z), V_m^{a+1}(x)\}u_f^b(y)dG(z)dy \\ &\quad + \left(1-dt\sum_{b=1}^{T}\int \lambda^{a+1,b}(x,y)dy\right)V_m^{a+1}(x)\Bigg]\Bigg\}.\end{aligned}$$

As all higher-order dt terms will vanish when taking $dt \to 0$, they can be ignored. After a bit of algebra, this equation simplifies to

$$\begin{aligned}(r+\rho+\psi_m^a(x))V_m^a(x) = (1-\psi_m^a(x)\,dt)\Bigg[&\rho V_m^{a+1}(x) \\ &+\sum_{b=1}^{T}\iint \lambda^{a,b}(x,y)\max\{S_m^{a,b}(x \mid y,z), 0\}u_f^b(y)dG(z)dy\Bigg].\end{aligned}$$

Taking $dt \to 0$ yields Equation (3.3.1).

C.2 Value function for married people

For a man married to a woman of type y with love shock z,

$$
\begin{aligned}
(1+r\,dt)W_m^{a,b}(x\mid y,z) = (1+r\,dt)t_m(x\mid y,z)\,dt + (1-\psi_m^a(x)\,dt)\Big[\psi_f^b(y)\,dt\times \\
\big((1-\rho\,dt)V_m^a(x)+\rho\,dt V_m^{a+1}(x)\big) + (1-\psi_f^b(y)\,dt)\times \\
\Big((1-\rho\,dt)\delta\,dt\int \max\{W_m^{a,b}(x\mid y,z), V_m^a(x)\}dG(z) \\
+\rho\,dt\delta\,dt\int \max\{W_m^{a+1,b+1}(x\mid y,z), V_m^{a+1}(x)\}dG(z) \\
+(1-\rho\,dt)(1-\delta\,dt)W_m^{a,b}(x\mid y,z)+\rho\,dt(1-\delta\,dt)W_m^{a+1,b+1}(x\mid y,z)\Big)\Big].
\end{aligned}
$$

Ignoring the higher-order dt terms, this equation simplifies to

$$
\begin{aligned}
(r+\rho+\psi_m^a(x)+\psi_f^b(y))W_m^{a,b}(x\mid y,z) = (1+r\,dt)t_m(x\mid y,z) \\
+(1-\psi_m^a(x)\,dt)\psi_f^b(y)V_m^a(x)+(1-\psi_m^a(x)\,dt)(1-\psi_f^b(y)\,dt)\times \\
\left[\rho(1-\delta\,dt)W_m^{a+1,b+1}(x\mid y,z)+(1-\rho\,dt)\delta\left(\int \max\{S_m^{a,b}(x\mid y,z),0\}dG(z)+V_m^a(x)\right)\right].
\end{aligned}
$$

Taking $dt\to 0$ yields

$$
\begin{aligned}
(r+\rho+\delta+\psi_m^a(x)+\psi_f^b(y))W_m^{a,b}(x\mid y,z) = t_m(x\mid y,z)+\psi_f^b(y)V_m^a(x) \\
+\rho W_m^{a+1,b+1}(x\mid y,z)+\delta\left(\int \max\{S_m^{a,b}(x\mid y,z),0\}dG(z)+V_m^a(x)\right).
\end{aligned}
$$

Rearranging gives Equation (3.3.2).

D Model without aging process

For clarity of exposition, consider the model of Goussé (2014) upon which my aging model is based. I use notation consistent with the main text, which differs slightly from the paper.

A steady state with no population growth implies

$$\ell_m(x) = \frac{\gamma_m(x)}{\psi_m(x)}. \tag{D.1}$$

D.1 Strategies

Denote the present value for a single man of type x by $V_m(x)$ and for a man of type x married to a woman of type y with love draw z by $W_m(x \mid y, z)$. Denote $S_m(x \mid y, z) := W_m(x \mid y, z) - V_m(x)$, the personal surplus. Then the male value functions are given by

$$(r + \psi_m(x))V_m(x) = \lambda \iint \max\{S_m(x \mid y, z), 0\}\, u_f(y) dG(z) dy \tag{D.2}$$

$$\begin{aligned} (r + \psi_m(x))W_m(x \mid y, z) = t_m(x \mid y, z) - (\delta + \psi_f(y))S_m(x \mid y, z) \\ + \delta \int \max\{S_m(x \mid y, z), 0\} dG(z). \end{aligned} \tag{D.3}$$

Denote the total match surplus by $\mathcal{S}(x, y, z) := S_m(x \mid y, z) + S_f(y \mid x, z)$. With Nash Bargaining, if females have bargaining power β, then $S_f(x \mid y, z) = \beta S(x, y, z)$. Since utility is transferable, the matching decision is always mutual: a pair marries if and only if $\mathcal{S}(x, y, z) > 0$.

Adding up value functions for a married couple and solving for the total surplus gives

$$\begin{aligned} (r + \delta + \psi_m(x) + \psi_f(y))\mathcal{S}(x, y, z) = z + f(x, y) + \delta \int \max\{\mathcal{S}(x, y, z), 0\}\, dG(z) \\ -(r + \psi_m(x))V_m(x) - (r + \psi_f(y))V_f(y). \end{aligned} \tag{D.4}$$

Thus, $\mathcal{S}(x, y, z)$ is linear in z and can be written as

$$(r + \delta + \psi_m(x) + \psi_f(y))\mathcal{S}(x, y, z) = s(x, y) + z$$

for some function s. Let $\alpha(x, y)$ be the probability of marriage conditional on meeting.

$$\begin{aligned}\alpha(x, y) &= \mathbb{P}[s(x, y) + z > 0] \\ &= 1 - \mathbb{P}[z \leq -s(x, y)] \\ &= 1 - G(-s(x, y)). \end{aligned} \tag{D.5}$$

D.2 Equilibrium

A steady state equilibrium requires that flows into and out of marriage are equal for each type of couple:

$$\forall x, y, \ (\psi_m(x) + \psi_f(y) + \delta(1 - \alpha(x, y)))m(x, y) = \lambda\alpha(x, y)u_m(x)u_f(y). \tag{D.6}$$

A steady state equilibrium is an assignment function $\alpha(x, y)$ (representing strategies), value functions $V_m(x)$, $V_f(y)$, and singles measures $u_m(x)$, $u_f(y)$ that solve Equations (D.2) to (D.6).

D.3 Estimation

With all the arrival rate parameters λ, δ in hand, α is non-parametrically identified by the steady state equilibrium condition in Equation (D.6):

$$\alpha(x, y) = \frac{(\psi_m(x) + \psi_f(y) + \delta)m(x, y)}{\lambda u_m(x)u_f(y) + \delta m(x, y)}.$$

As the scale of $s(x, y)$ is not identified, the distribution of love shocks, G, can be normalized to fix the scale. Then $s(x, y)$ is recovered by inverting Equation (D.5):

$$s(x, y) = -G^{-1}(1 - \alpha(x, y)).$$

Finally, to recover $f(x, y)$, it is necessary to first compute the singlehood value functions.

This requires solving the integral in the value functions in Equation (D.2):

$$
\begin{aligned}
\int \max\{\mathcal{S}(x,y,z),0\}\,dG(z) &= \frac{\int_{-s(x,y)}^{\infty}[s(x,y)+z]\,dG(z)}{r+\delta+\psi_m(x)+\psi_f(y)} \\
&= \frac{(1-G(-s(x,y)))(s(x,y)+\mathbb{E}[z \mid z>-s(x,y)])}{r+\delta+\psi_m(x)+\psi_f(y)} \\
&= \frac{\alpha(x,y)(s(x,y)+\mathbb{E}[z \mid z>-s(x,y)])}{r+\delta+\psi_m(x)+\psi_f(y)} \\
&:= \frac{\mu(\alpha(x,y))}{r+\delta+\psi_m(x)+\psi_f(y)}.
\end{aligned}
$$

Then the singlehood value functions reduce to

$$
(r+\psi_m(x))V_m(x) = (1-\beta)\lambda \int \frac{\mu(\alpha(x,y))}{r+\delta+\psi_m(x)+\psi_f(y)} u_f(y)dy. \tag{D.7}
$$

Now $f(x,y)$ can be recovered from Equation (D.4) using $s(x,y)$ along with the linearity result.

D.4 Computation of equilibrium

Integrating Equation (D.6) over the opposite sex yields the steady state singles distribution:

$$
\ell_m(x) - u_m(x) = \int m(x,y)dy = \lambda u_m(x) \int \frac{\alpha(x,y)}{\psi_m+\psi_f+\delta(1-\alpha(x,y))} u_f(y)dy \tag{D.8}
$$

$$
\iff u_m(x) = \frac{\ell_m(x)}{1+\lambda\int \frac{\alpha(x,y)}{\psi_m+\psi_f+\delta(1-\alpha(x,y))}u_f(y)dy}. \tag{D.9}
$$

Using the above derivation, I can use $s(x,y) = -G^{-1}(1-\alpha(x,y))$ to write the integral only in terms of α:

$$
\int \max\{\mathcal{S}(x,y,z),0\}\,dG(z) = \frac{\mu(\alpha(x,y))}{r+\delta+\psi_m(x)+\psi_f(y)}
$$

Substituting into Equation (D.4) and using the mapping in Equation (D.5), the conditional matching probability solves

$$\begin{aligned}\alpha(x,y) = 1 - G\bigg(&- f(x,y) - \frac{\delta}{r+\delta+\psi_m(x)+\psi_f(y)}\mu(\alpha(x,y)) \\ &+(r+\psi_m(x))V_m(x) + (r+\psi_f(y))V_f(y)\bigg).\end{aligned} \tag{D.10}$$

An equilibrium can be computed as a fixed point of the following mapping:

1. Start with an initial guess for α.
2. Use α to update u_m, u_f from the steady state flow condition in Equation (D.8).
3. Solve Equation (D.7) for implied value functions.
4. Use the value functions to compute the implied α from Equation (D.10).

E Continuous aging model

Denote the type spaces by $X = [\underline{x}, \overline{x}]$ and $Y = [\underline{y}, \overline{y}]$. Agent types depreciate at rate ρ. Treating types as "youth", assume that young adults of types $\overline{x}$ and $\overline{y}$ enter the marriage market at rate γ_0 (and there is no other entry).

E.1 Steady state population distribution

The steady state population is given by equating inflows with outflows. Consider a small time increment dt and the subpopulation $[x, \overline{x}]$. The outflows from this subpopulation are from death and aging, while the inflows are from birth.

$$\begin{aligned}
dt \int_x^{\overline{x}} \psi_m(v)\ell_m(v)dv + \int_x^{x(1+\rho dt)} \ell_m(v)dv &= \gamma_0 dt \\
\iff \int_x^{\overline{x}} \psi_m(v)\ell_m(v)dv + \rho x \frac{L_m(x+\rho x dt) - L_m(x)}{\rho x dt} &= \gamma_0 \\
dt \to 0 \implies \int_x^{\overline{x}} \psi_m(v)\ell_m(v)dv + \rho x \ell_m(x) &= \gamma_0.
\end{aligned}$$

As this holds for every x, this equation can be differentiated:

$$\begin{aligned}
\rho\ell_m(x) + \rho x \ell_m'(x) &= \psi_m(x)\ell_m(x) \\
\iff \frac{\ell_m'(x)}{\ell_m(x)} &= \frac{\psi_m(x) - \rho}{\rho x}.
\end{aligned}$$

This ordinary differential equation has a solution of the form

$$\begin{aligned}
\log(\ell_m(x)) &= \frac{1}{\rho}\int \frac{\psi(x)}{x} - \log(x) + c_0 \\
\ell_m(x) &= \frac{c_1}{x} \exp\left(\frac{1}{\rho}\int \frac{\psi(x)}{x}\right).
\end{aligned}$$

Using the boundary condition $\ell_m(\overline{x}) = \gamma_0$ pins down the constant term:

$$c_1 = \frac{\overline{x}\gamma_0}{\exp\left(\frac{1}{\rho}\int \frac{\psi(\overline{x})}{\overline{x}}\right)}.$$

Thus, the solution gives the steady state population distribution,

$$\begin{aligned}\ell_m(x) &= \gamma_0 \frac{\overline{x}}{x} \exp\left(\int \frac{\psi(x)}{\rho x} - \int \frac{\psi(\overline{x})}{\rho \overline{x}} \right) \\ &= \frac{\gamma_0}{\frac{x}{\overline{x}} \exp\left(\frac{1}{\rho} \int_x^{\overline{x}} \frac{\psi(v)}{v} dv \right)}.\end{aligned}$$

E.2 Strategies

Value functions are derived similarly, but they now contain a partial derivative term:

$$(r + \psi_m(x))V_m(x) + x\rho \frac{\partial V_m}{\partial x}(x) = \lambda \iint \max\left\{S_m(x \mid y, z), 0\right\} u_f(y) dG(z) dy \quad \text{(E.1)}$$

$$\begin{aligned}(r + \psi_m(x))W_m(x \mid y, z) + \rho \left(x \frac{\partial W_m}{\partial x}(x \mid y, z) + y \frac{\partial W_m}{\partial y}(x \mid y, z) \right) &= t_m(x \mid y, z) - \quad \text{(E.2)} \\ (\delta + \psi_f(y))S_m(x \mid y, z) + \delta \int \max\{S_m(x \mid y, z), 0\} dG(z).\end{aligned}$$

The surplus, however, is no longer easily solvable. Adding up value functions for a married couple and solving for the total surplus gives

$$\begin{aligned}(r + \delta + \psi_m(x) + \psi_f(y))\mathcal{S}(x, y, z) + \rho \left(x \frac{\partial \mathcal{S}}{\partial x}(x, y, z) + y \frac{\partial \mathcal{S}}{\partial y}(x, y, z) \right) &= z + \\ f(x, y) + \delta \int \max\left\{\mathcal{S}(x, y, z), 0\right\} dG(z) - \quad \text{(E.3)} \\ (r + \psi_m(x))V_m(x) - (r + \psi_f(y))V_f(y) - \rho \left(x \frac{\partial V_m}{\partial x}(x) + y \frac{\partial V_f}{\partial y}(y) \right).\end{aligned}$$

Using the method of "guess and check", it can be shown that $\mathcal{S}$ is linear in z.

E.3 Equilibrium

The steady-state flow balance condition for marriages is now given by a partial differential equation, which accounts for aging flows:

$$m(x, y) \left(\psi_m(x) + \psi_f(y) + \delta(1 - \alpha(x, y) + 2\rho)\right) = \lambda u_m(x) u_f(y) \alpha(x, y) - \rho \nabla m(x, y) \quad \text{(E.4)}$$

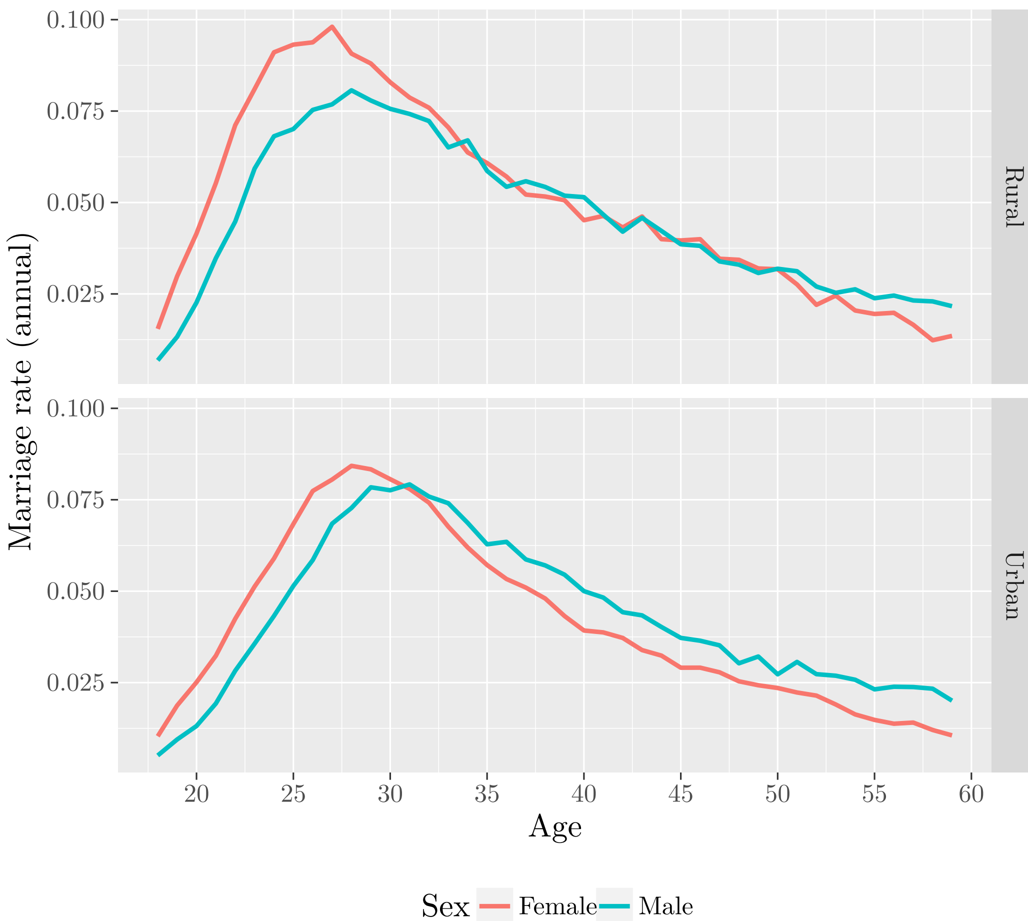
This is numerically expensive to solve.

F Supplemental tables and figures

Rank	Population	MSA
1	14.5	New York-Newark-Jersey City, NY-NJ-PA
2	9.4	Los Angeles-Long Beach-Anaheim, CA
3	6.8	Chicago-Naperville-Elgin, IL-IN-WI
4	4.6	Dallas-Fort Worth-Arlington, TX
5	4.4	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
6	4.2	Houston-The Woodlands-Sugar Land, TX
7	4.1	Washington-Arlington-Alexandria, DC-VA-MD-WV
8	4.1	Miami-Fort Lauderdale-West Palm Beach, FL
9	3.8	Atlanta-Sandy Springs-Roswell, GA
10	3.5	Boston-Cambridge-Newton, MA-NH
11	3.3	San Francisco-Oakland-Hayward, CA
12	3.1	Detroit-Warren-Dearborn, MI
13	3.1	Phoenix-Mesa-Scottsdale, AZ
14	3.0	Riverside-San Bernardino-Ontario, CA
15	2.6	Seattle-Tacoma-Bellevue, WA
16	2.4	Minneapolis-St. Paul-Bloomington, MN-WI
17	2.3	San Diego-Carlsbad, CA
18	2.1	Tampa-St. Petersburg-Clearwater, FL
19	2.0	St. Louis, MO-IL
20	2.0	Baltimore-Columbia-Towson, MD

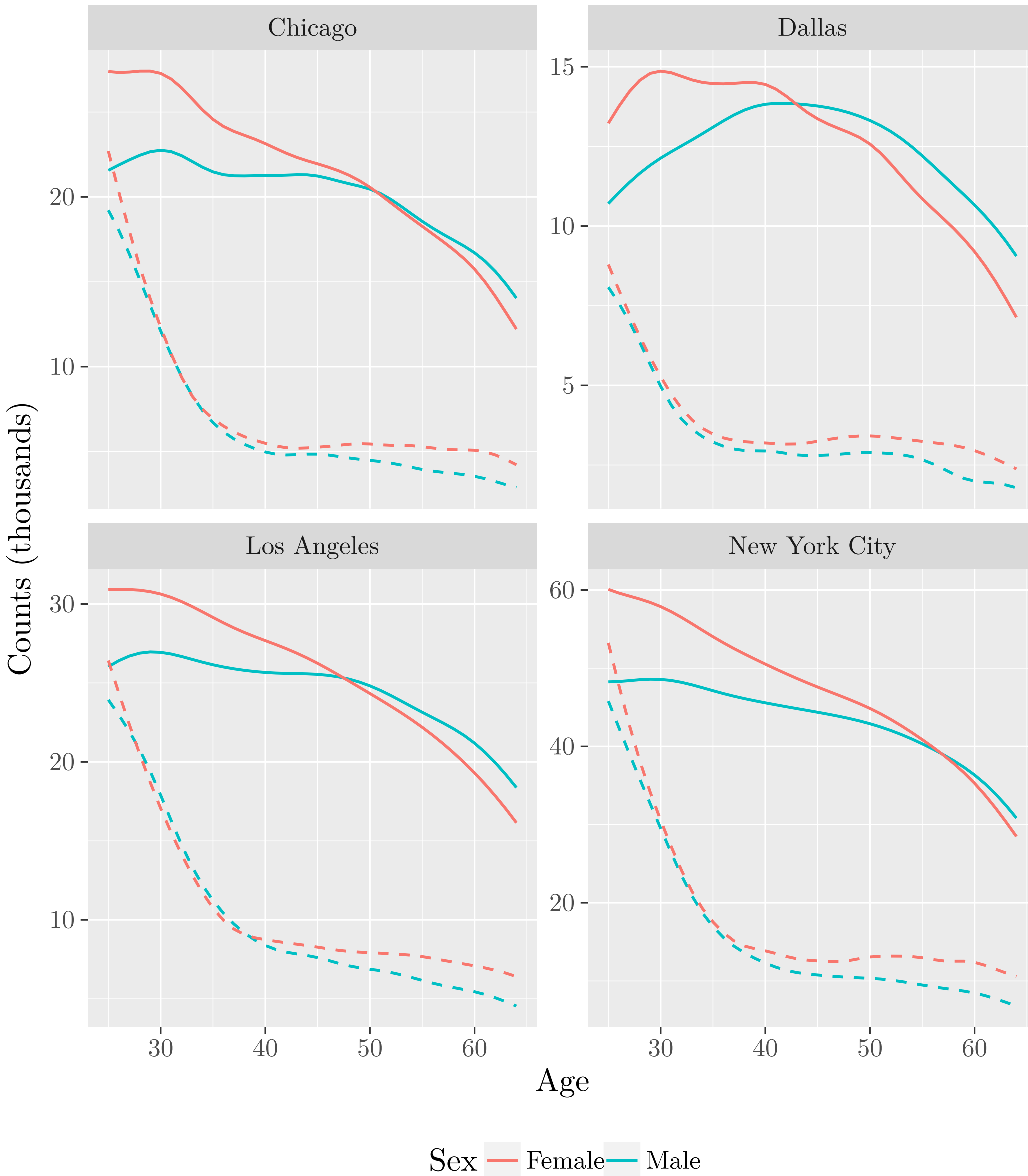
Source: ACS 2014.

Table F.1: Top 20 MSAs by population (in millions)



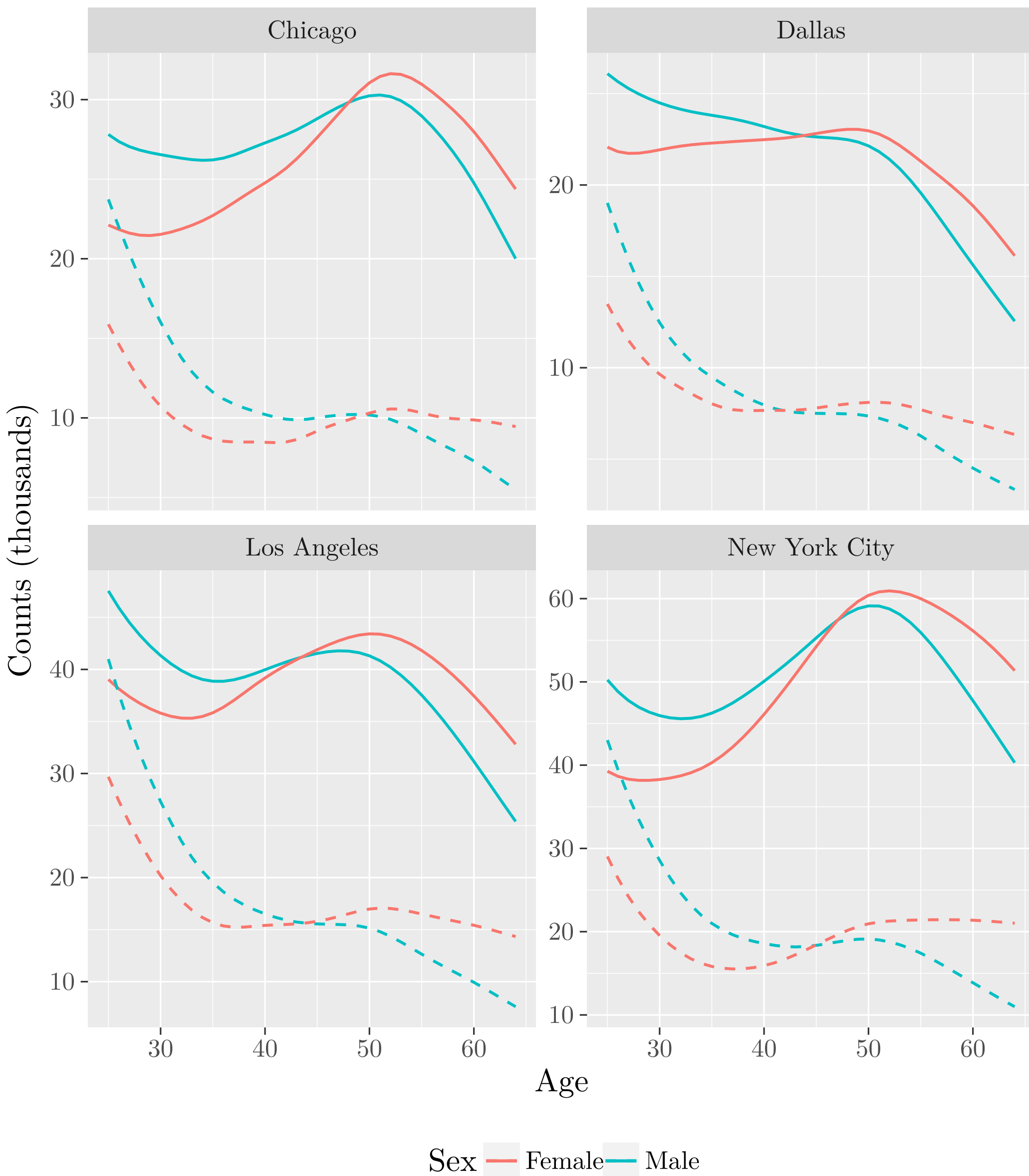
Notes: The marriage rate is the proportion of all singles who enter into marriage during the survey year. ACS 2008–2016.

Figure F.1: Marriage rates by sex and rural vs urban status



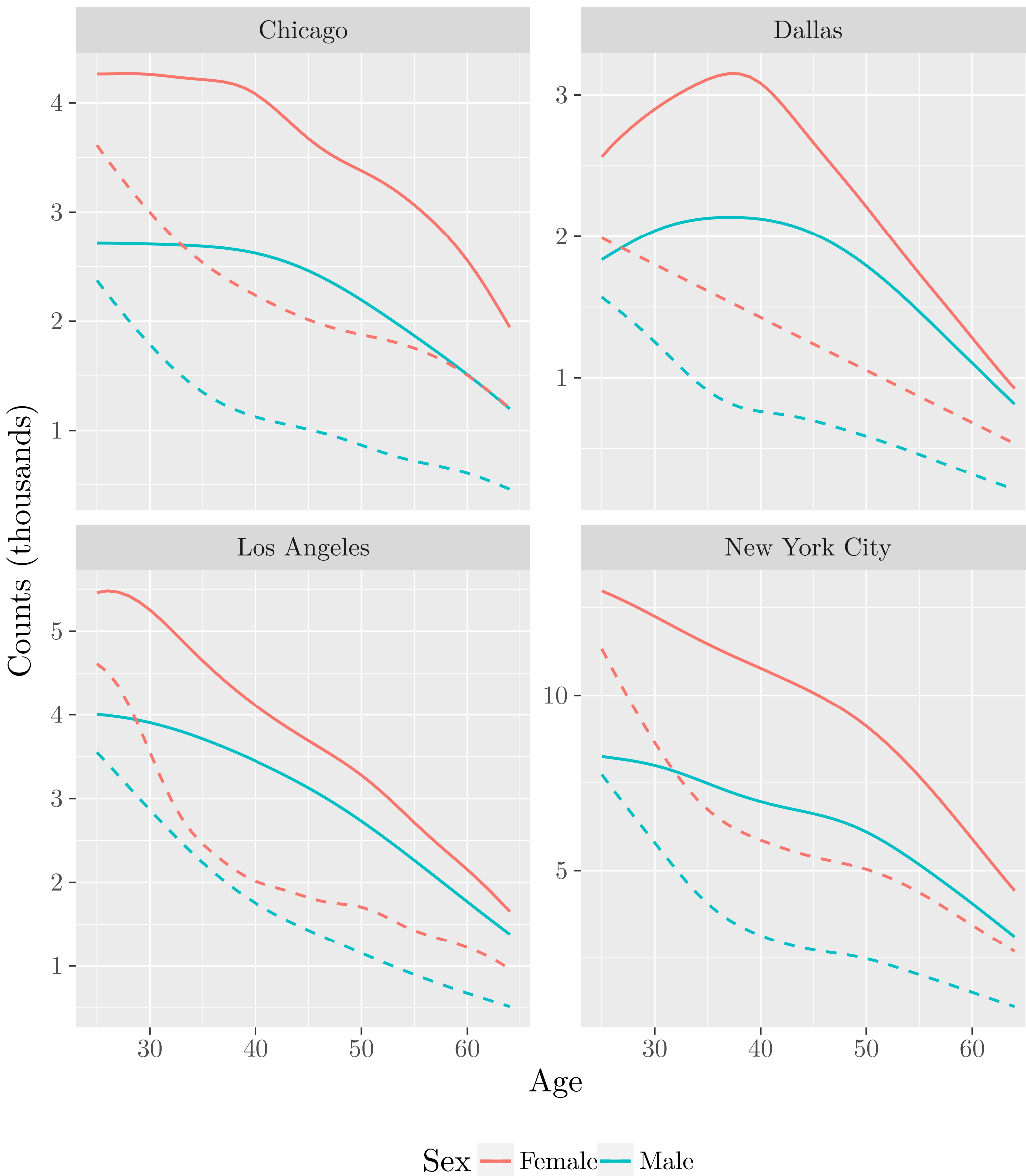
Notes: Counts are smoothed as described in Section 4.2. The solid lines denote total population counts whereas the dotted lines denote counts of singles. The y-axis is cropped from below for clarity. ACS 2008–2016.

Figure F.2: Populations of college-educated non-minority couples



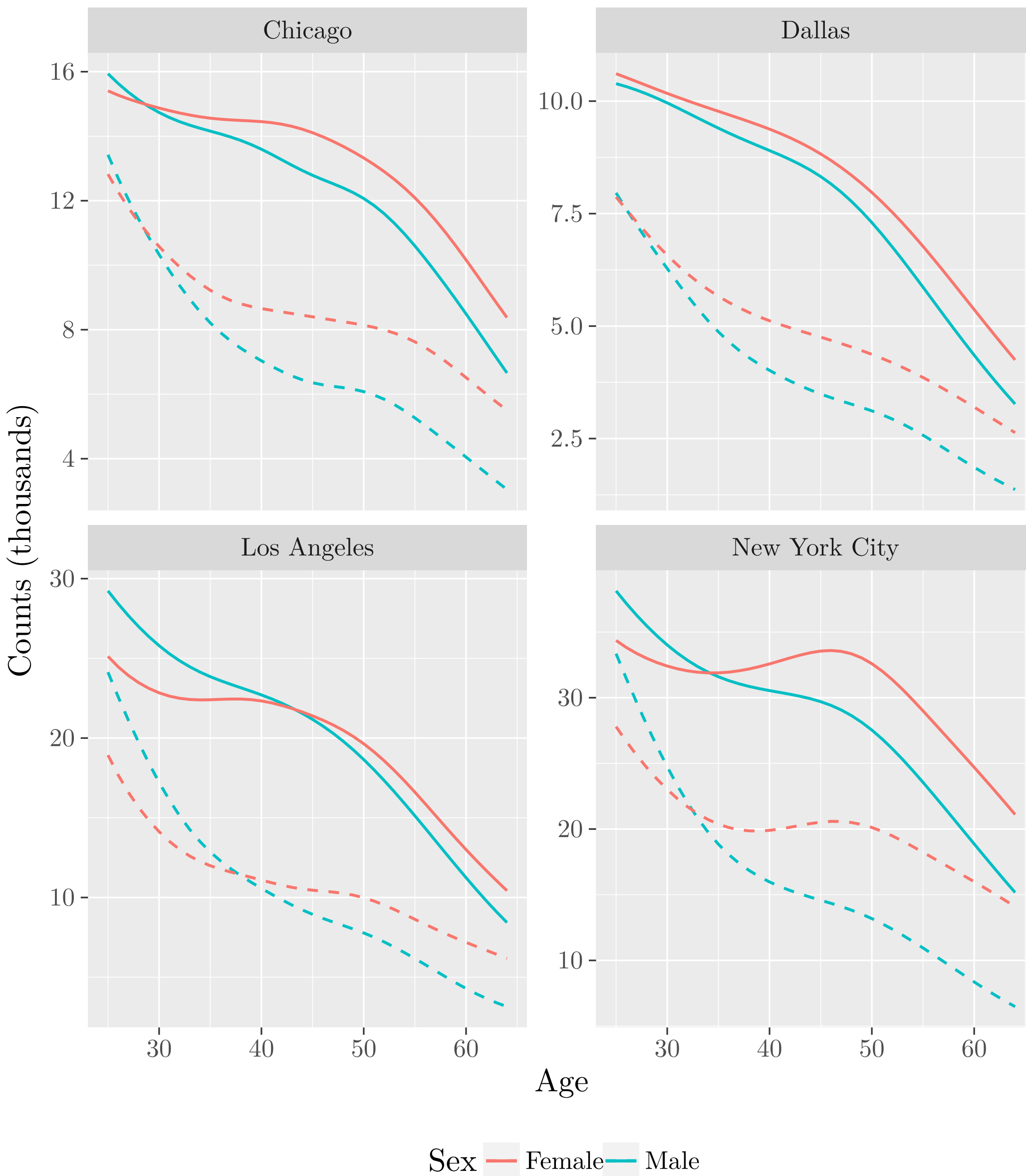
Notes: Counts are smoothed as described in Section 4.2. The solid lines denote total population counts whereas the dotted lines denote counts of singles. The y-axis is cropped from below for clarity. ACS 2008–2016.

Figure F.3: Populations of non-college non-minority couples



Notes: Counts are smoothed as described in Section 4.2. The solid lines denote total population counts whereas the dotted lines denote counts of singles. The y-axis is cropped from below for clarity. ACS 2008–2016.

Figure F.4: Populations of college-educated minority couples



Notes: Counts are smoothed as described in Section 4.2. The solid lines denote total population counts whereas the dotted lines denote counts of singles. The y-axis is cropped from below for clarity. ACS 2008–2016.

Figure F.5: Populations of non-college minority couples

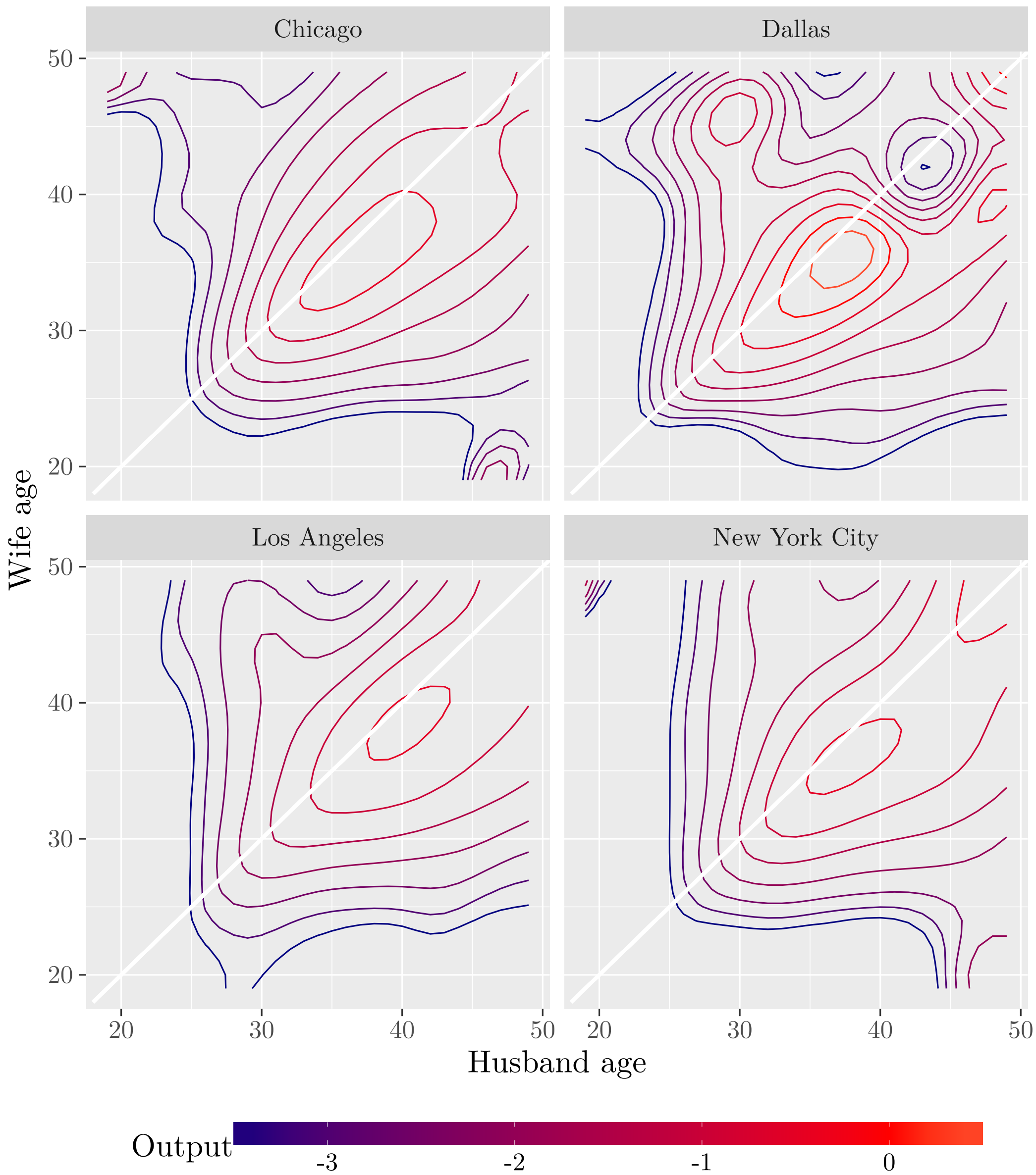


Figure F.6: Marital flow output

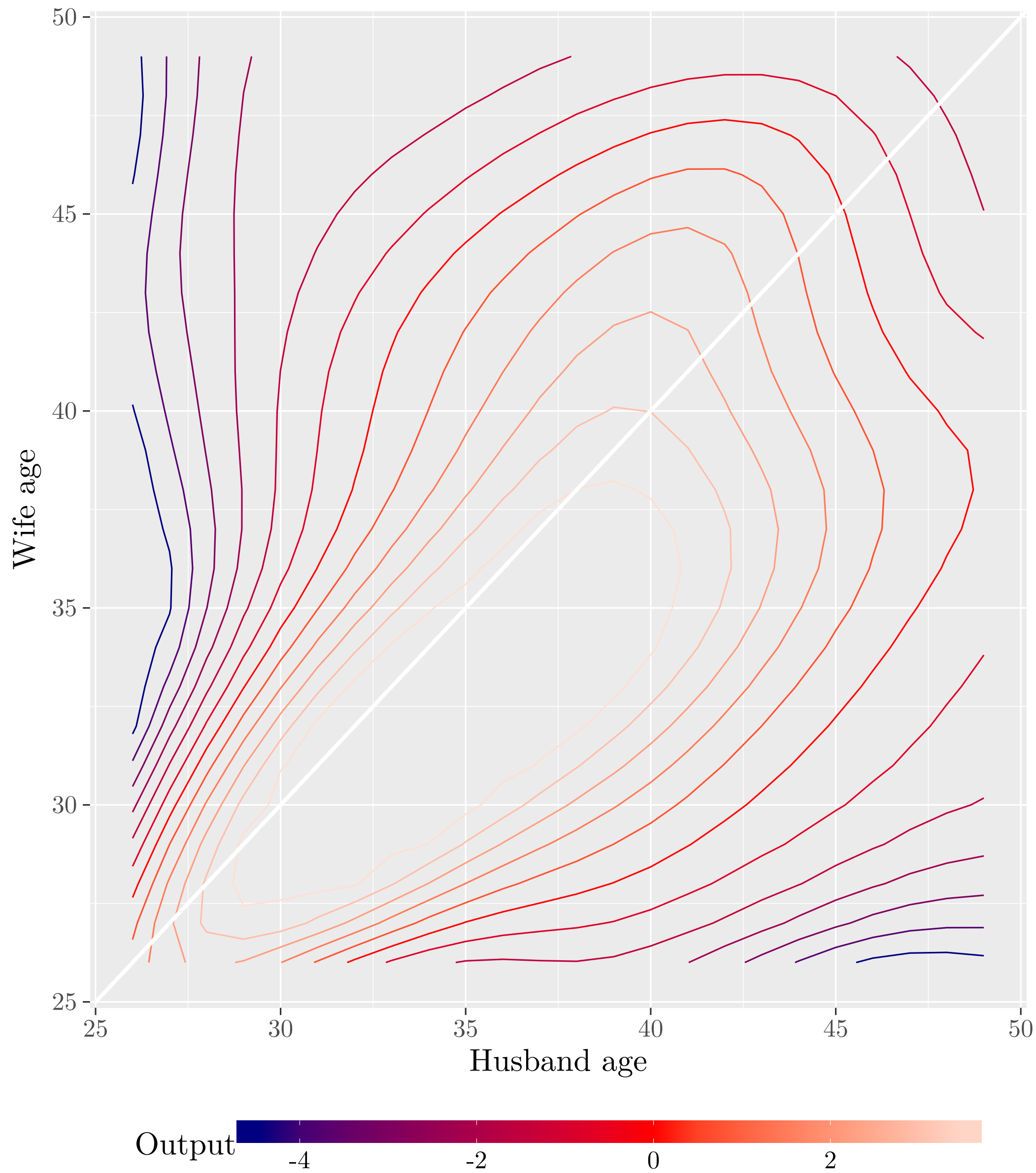


Figure F.7: Marital flow output for college-educated non-minority couples

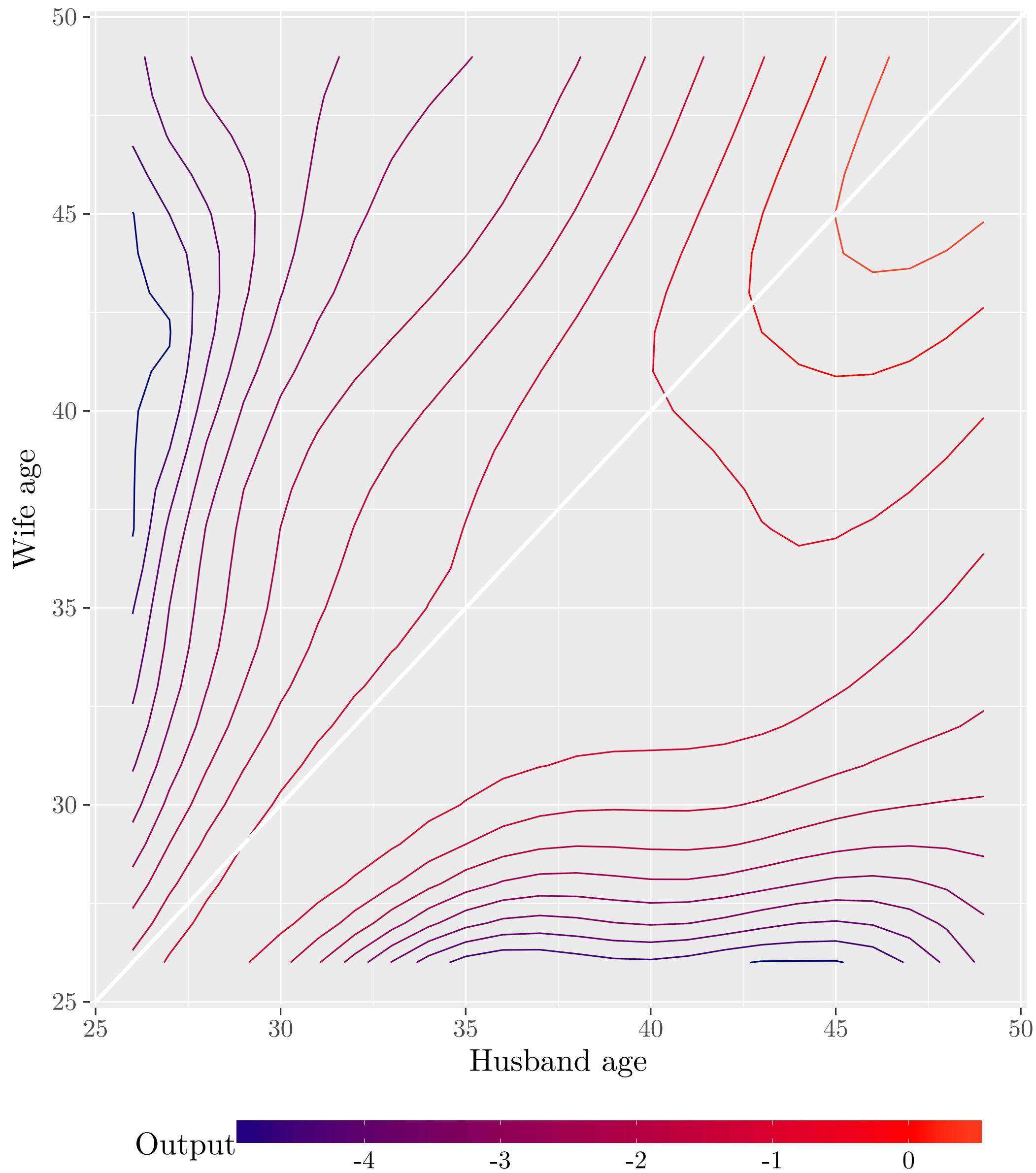


Figure F.8: Marital flow output for non-college non-minority couples

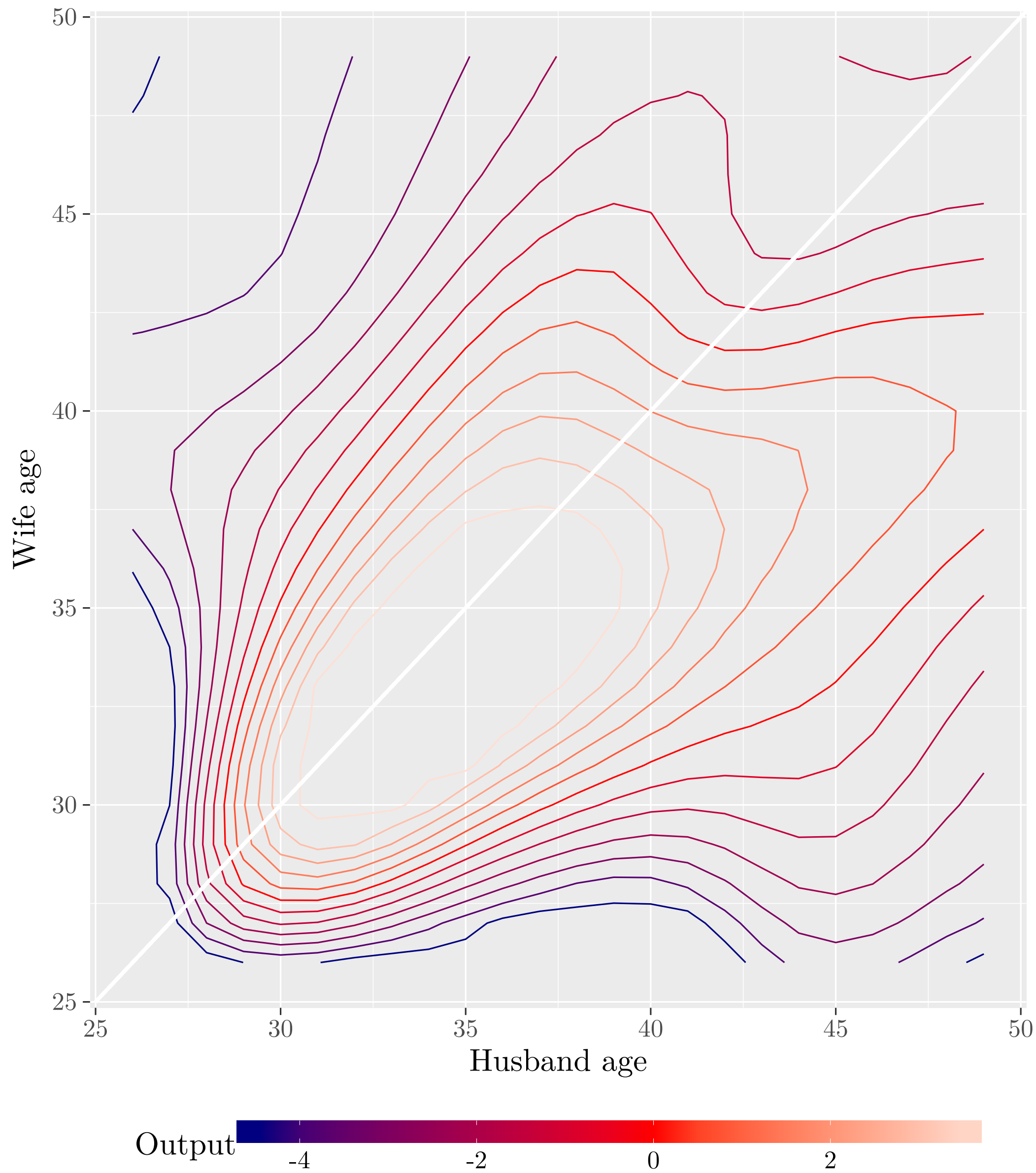


Figure F.9: Marital flow output for college-educated minority couples

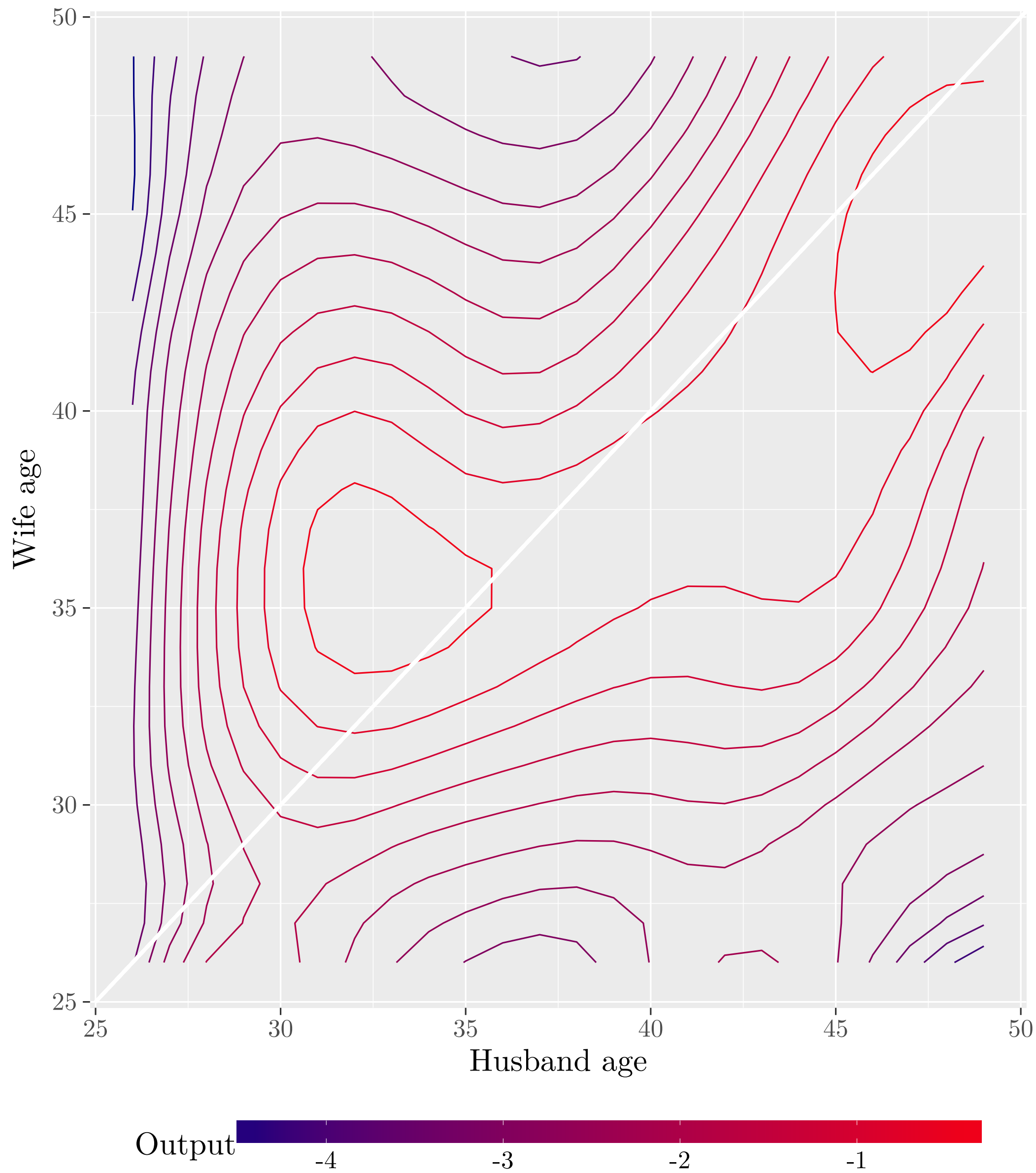


Figure F.10: Marital flow output for non-college minority couples

Figure F.11: Value of singlehood for college-educated non-minority couples

Figure F.12: Value of singlehood for non-college non-minority couples

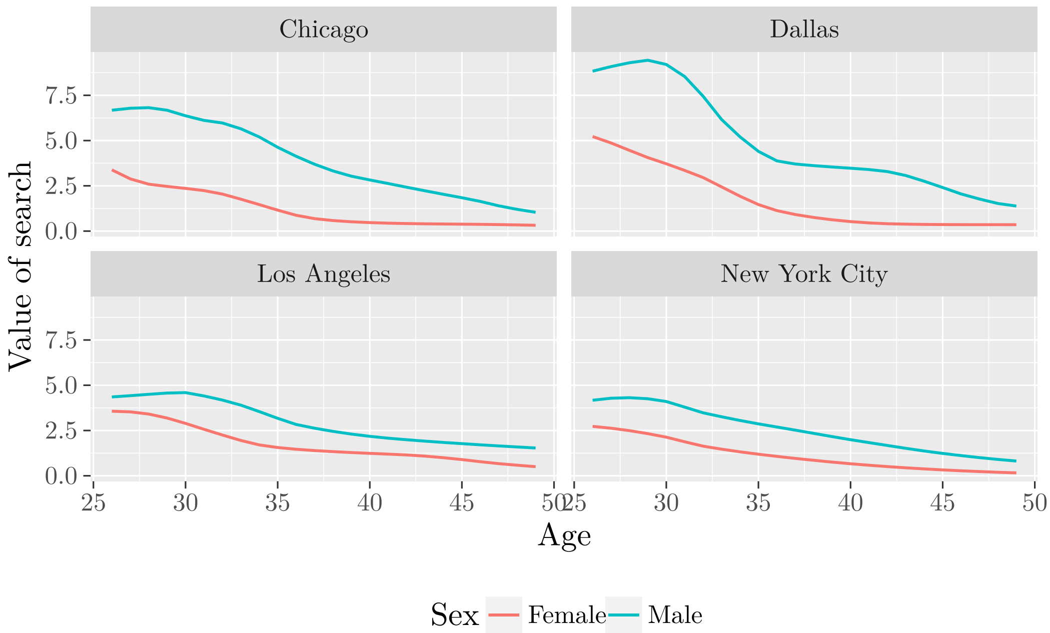


Figure F.13: Value of singlehood for college-educated minority couples

Figure F.14: Value of singlehood for non-college minority couples

Bibliography

"A distorted sex ratio is playing havoc with marriage in China" (Nov. 2017). In: *The Economist* (Special report: A looser knot).

Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos (2011). "Marrying up: the role of sex ratio in assortative matching". In: *American Economic Journal: Applied Economics*, pp. 124–157.

Altonji, Joseph G and Lewis M Segal (1996). "Small-sample bias in GMM estimation of covariance structures". In: *Journal of Business & Economic Statistics* 14.3, pp. 353–366.

Angrist, Josh (2002). "How do sex ratios affect marriage and labor markets? Evidence from America's second generation". In: *Quarterly Journal of Economics*, pp. 997–1038.

Autor, David, David Dorn, Gordon Hanson, et al. (2017). "When work disappears: manufacturing decline and the falling marriage-market value of men". NBER Working Paper.

Becker, Gary S (1973). "A theory of marriage: Part I". In: *Journal of Political Economy* 81.4, pp. 813–846.

Birger, Jon (2015). *Date-onomics*. Workman Publishing Company.

Brien, Michael J, Lee A Lillard, and Steven Stern (2006). "Cohabitation, marriage, and divorce in a model of match quality". In: *International Economic Review* 47.2, pp. 451–494.

Bronson, Mary Ann (2015). "Degrees are forever: Marriage, educational investment, and lifecycle labor decisions of men and women". Working Paper.

Bronson, Mary Ann and Maurizio Mazzocco (2018). "Cohort size and the marriage market: What explains the negative relationship?" Working Paper.

Burdett, Ken and Melvyn G. Coles (1997). “Marriage and class”. In: *The Quarterly Journal of Economics*, pp. 141–168.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss (2017). “Partner choice, investment in children, and the marital college premium”. In: *American Economic Review* 107.8, pp. 2109–67.

Choo, Eugene (2015). “Dynamic marriage matching: An empirical framework”. In: *Econometrica* 83.4, pp. 1373–1423.

Choo, Eugene and Aloysius Siow (2006). “Who marries whom and why”. In: *Journal of political Economy* 114.1, pp. 175–201.

Coles, Melvyn G. and Marco Francesconi (2011). “On the emergence of toyboys: The timing of marriage with aging and uncertain careers”. In: *International Economic Review* 52.3, pp. 825–853.

— (2017). “Equilibrium Search and the Impact of Equal Opportunities for Women”. SSRN.

Dessy, Sylvain and Habiba Djebbari (2010). “High-powered careers and marriage: can women have it all?”

Díaz-Giménez, Javier and Eugenio Giolito (2013). “Accounting for the Timing of First Marriage”. In: *International Economic Review* 54.1, pp. 135–158.

Edlund, Lena (2013). “The role of paternity presumption and custodial rights for understanding marriage patterns”. In: *Economica* 80.320, pp. 650–669.

Flabbi, Luca and Christopher Flinn (2015). “Simultaneous Search in the Labor and Marriage Markets with Endogenous Schooling Decisions”. Unpublished manuscript.

Gale, David and Lloyd S Shapley (1962). “College admissions and the stability of marriage”. In: *The American Mathematical Monthly* 69.1, pp. 9–15.

Galichon, Alfred and Bernard Salanié (2015). “Cupid’s invisible hand: Social surplus and identification in matching models”. SSRN.

Gayle, George-Levi and Andrew Shephard (2018). “Optimal taxation, marriage, home production, and family labor supply”. In: *Econometrica (Forthcoming)*.

Goussé, Marion (2014). “Marriage gains: Who should you marry?” SSRN.

Goussé, Marion, Nicolas Jacquemet, and Jean-Marc Robin (2017a). “Household Labour Supply and the Marriage Market in the UK, 1991-2008”. In: *Labour Economics* 46, pp. 131–149.

— (2017b). “Marriage, Labor Supply, and Home Production”. In: *Econometrica* 85.6, pp. 1873–1919.

Greenwood, Jeremy et al. (2014). “Marry your like: Assortative mating and income inequality”. In: *American Economic Review* 104.5, pp. 348–53.

Guttentag, Marcia and Paul F Secord (1983). *Too many women? The sex ratio question.* London England Sage Publications.

Johnson, Steven G (2014). *The NLopt nonlinear-optimization package.*

Li, Qi and Jeffrey Scott Racine (2007). *Nonparametric econometrics: theory and practice.* Princeton University Press.

Low, Corinne (2015). “Pricing the Biological Clock: Reproductive Capital on the US Marriage Market”. Working Paper.

Lundberg, Shelly and Robert A Pollak (2007). “The American Family and Family Economics”. In: *Journal of Economic Perspectives* 21.2, pp. 3–26.

— (2013). “Cohabitation and the Uneven Retreat from Marriage in the US, 1950-2010”. NBER Working Paper.

Powell, MJD (1998). “Direct search algorithms for optimization calculations”. In: *Acta numerica* 7, pp. 287–336.

Rıos-Rull, José-Vıctor, Shannon Seitz, and Satoshi Tanaka (2016). “Sex Ratios and Long-Term Marriage Trends”. Working Paper.

Ruggles, Steven et al. (2017). *Integrated Public Use Microdata Series: Version 7.0 [dataset].* University of Minnesota.

Seitz, Shannon (2009). “Accounting for racial differences in marriage and employment”. In: *Journal of Labor Economics* 27.3, pp. 385–437.

Shimer, Robert and Lones Smith (2000). “Assortative matching and search”. In: *Econometrica* 68.2, pp. 343–369.

Shin, Seul-ki (2014). “Preferences vs. Opportunities: Racial/Ethnic Intermarriage in the United States”. PIER Working Paper.

Siow, Aloysius (1998). “Differential fecundity, markets, and gender roles”. In: *Journal of Political Economy* 106.2, pp. 334–354.

Tange, Ole (Apr. 2018). *GNU Parallel 2018*. Ole Tange.

Trivers, Robert L (1972). “Parental investment and sexual selection”. In: *Sexual Selection & the Descent of Man*. Aldine de Gruyter, pp. 136–179.

Wide-ranging OnLine Data for Epidemiologic Research (WONDER) (2017). Centers for Disease Control and Prevention. URL: `http://wonder.cdc.gov`.

Wilson, William Julius (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

Wong, Linda Y (2003a). “Structural estimation of marriage models”. In: *Journal of Labor Economics* 21.3, pp. 699–727.

— (2003b). “Why do only 5.5% of Black Men Marry White Women?” In: *International Economic Review* 44.3, pp. 803–826.