



HAL
open science

Matching models with and without frictions: applications to the economics of the family

Edoardo Ciscato

► **To cite this version:**

Edoardo Ciscato. Matching models with and without frictions: applications to the economics of the family. Economics and Finance. Institut d'études politiques de paris - Sciences Po, 2019. English. NNT : 2019IEPP0018 . tel-03411916

HAL Id: tel-03411916

<https://theses.hal.science/tel-03411916>

Submitted on 14 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Institut d'études politiques de Paris
ÉCOLE DOCTORALE DE SCIENCES PO
Programme doctoral en économie
Département d'économie
Doctorat en Sciences économiques

Matching models with and without frictions: Applications to the economics of the family

Edoardo Ciscato

Thesis supervised by:

Jean-Marc Robin, Professeur des Universités, Sciences Po Paris

Defended in Paris on June 20, 2019

Jury:

Mr. Pierre-André Chiappori, Professor of Economics, Columbia University

Mr. Arnaud Dupuy, Full Professor of Economics, University of Luxembourg (*reviewer*)

Mr. Alfred Galichon, Professor of Economics, New-York University

Mrs. Marion Goussé, Professeur agrégée, Université Laval

Mr. Frederic Vermeulen, Professor of Economics, University of Leuven (*reviewer*)

Abstract

This dissertation aims to shed light on how the institution of marriage has evolved in the United States over the last 60 years. In the first chapter, joint work with Alfred Galichon and Marion Goussé, we extend Gary Becker's equilibrium model of marriage and the family to analyze same-sex marriage and develop an econometric method to study sorting patterns among same-sex couples. We discuss differences in the gains from marriage between same-sex and different-sex couples using a sample of Californian households for the period from 2008 to 2012. In the second chapter, joint work with Simon Weber, we describe mating patterns in the United States from 1964 to 2017 and measure the impact of changes in marital preferences on between-household income inequality. We estimate the strength of positive assortative mating with respect to socio-economic characteristics for different cohorts and answer the following questions: has assortativeness increased over time? If yes, along which dimensions? And to what extent the shifts in marital preferences can explain inequality trends? In the third chapter, I build a novel equilibrium model of the marriage market characterized by search frictions, endogenous divorce, aging and wage mobility. I estimate the model with American data for two separate periods, the 1970s and the 2000s, and then provide a quantitative assessment of the impact of changes in the wage distribution on the decline of marriage observed between these two periods. I conclude by discussing the impact of changes in the wage distribution on the welfare of different population groups.

Contents

List of Figures	vii
List of Tables	ix
Acknowledgments	xi
Introduction	1
Bibliography	8
Chapter 1. Like attract like? A structural comparison of homogamy across same-sex and different-sex households ¹	11
1.1. Introduction	12
1.2. Model	16
1.3. Estimation	21
1.4. Data	27
1.5. Results	30
1.6. Discussion and perspectives	37
1.A. Pooled matching market with sexual orientation	39
1.B. Descriptive statistics	40
1.C. Main estimation results	41
1.D. Bipartite estimation for same-sex couples: head/spouse	43
1.E. Robustness checks	44
Bibliography	45
Chapter 2. The role of evolving marital preferences in growing income inequality ²	49
2.1. Introduction	49
2.2. Previous findings	52
2.3. Theoretical framework	56
2.4. Data	60
2.5. Trends in matching patterns	67
2.6. Counterfactual analysis	74
2.7. Conclusions and perspectives	80
2.A. Technical appendix	81
2.B. Samples	83
2.C. Additional Figures	83

¹Chapter coauthored with Alfred Galichon and Marion Goussé.

²Chapter coauthored with Simon Weber.

Bibliography	89
Chapter 3. The changing wage distribution and the decline of marriage	93
3.1. Introduction	93
3.2. Model	99
3.3. Identification	114
3.4. Data	118
3.5. Empirical specification and estimation	121
3.6. Results	127
3.7. Conclusion	134
3.A. Technical Appendix	135
3.B. Tables	138
3.C. Figures	147
Bibliography	159

List of Figures

2.1 Median age at first marriage	64
2.2 Share of married individuals by age and cohort	64
2.3 Summary Statistics	66
2.4 Partners' Traits	67
2.5 Assortativeness in age	68
2.6 Assortativeness in education	69
2.7 Assortativeness in wage	70
2.8 Assortativeness in Hours of work	71
2.9 Assortativeness in Race, by number of race included	72
2.10 Sigma	74
2.11 Assortativeness in education, counterfactual distribution	75
2.12 Assortativeness in age, counterfactual distribution	76
2.13 Counterfactual analysis, Gini coefficients since 1971	78
2.14 Decomposition of Gini coefficient shift 1971-2017 due to marital preferences	79
2.15 Relevant cross-interactions	84
2.16 Sample selection by median age	85
2.17 Sorting on annual earnings	85
2.18 All couples and working couples (where both partners work)	86
2.19 All couples and childless couples	86
2.20 Sorting on potential income	87
2.21 Decomposed taste for interracial marriage	87
2.22 Education omitted	88
2.23 Cohabiting and married couples	88
3.1 Timeline for a man i and a woman j married in t	147
3.2 Timeline for a woman j who is single in t	147
3.3 Wage levels by gender and age	148
3.4 College premium and gender wage gap	149
3.5 Changes in wage mobility by gender, age and education	150

3.6 Estimates of γ_2 by age and education	151
3.7 Fit of the model: share of married men by age and education	152
3.8 Fit of the model: share of married women by age and education	153
3.9 Fit of the model: changes on the extensive margin by gender and age	154
3.10 Fit of the model: changes on the extensive margin by gender and education	155
3.11 Changes in the share of married men by age and education: experiment (a)	156
3.12 Changes in intertemporal welfare of young individuals: experiment (a)	157
3.13 Changes in intertemporal welfare of young individuals: from the 1970s to the 2000s	158

List of Tables

1.1 Sample means by sexual orientation	41
1.2 Distribution of race by sexual orientation	41
1.3 Couples' Pearson correlation coefficients by sexual orientation	41
1.4 Homogamy rates by sexual orientation	41
1.5 Estimated affinity matrix by marriage market	42
1.6 Estimated indices of attractiveness by marriage market	43
1.7 Estimated affinity matrix for same-sex couples with asymmetric roles	44
1.8 Summary tables: main findings and robustness checks	44
2.1 Robustness Checks	83
3.1 Estimates of γ_1 by spouses' education	138
3.2 Estimates of the search parameters	139
3.3 Summary of parameters estimated or calibrated outside of the model	139
3.4 Fit of the model (1)	140
3.5 Fit of the model (2)	141
3.6 Decomposition of the changes in marriage market outcome (1)	142
3.7 Decomposition of the changes in marriage market outcome (2)	143
3.8 Changes in wage mobility and wage levels (1)	144
3.9 Changes in wage mobility and wage levels (2)	145
3.10 Changes in the distribution of intertemporal welfare at the beginning of adulthood	146

Acknowledgments

I would like to start by thanking Alfred Galichon, Jean-Marc Robin and Pierre-André Chiappori, who greatly inspired and guided me through my first steps in the world of economic research.

I would like to thank Alfred first, as he was the first that I met at the very beginning of my Master. I am very grateful to him for many reasons: he hired me as a Research Assistant, taught me about the econometrics of matching, invited me to NYU, introduced me to many of the people that are listed below and gave me the opportunity of working with him on a paper - the first chapter of this thesis - that proved to be successful.

I am also extremely grateful to Jean-Marc, who accepted to advise me during my PhD, helped me develop my Job Market Paper and patiently supported me during the Job Market. He showed me how to conduct rigorous empirical research, introduced me to search frictions and even taught me to discuss about economics in French.

Finally, I would like to thank Pierre-André: I have learned a lot from his papers, slides and handbooks, but his passion and personal feedback have been even more important to me. The months I spent at Columbia under his supervision were extremely productive and turned out to be a turning point during my PhD.

I am also deeply indebted to two former Sciences Po PhD students and by now successful researchers, Marion Goussé and Simon Weber, who coauthored with me the first and second chapter of this thesis respectively. I have learned a lot while working with them, starting from the rudiments of LaTeX, going through some more serious coding and finally to the joys and pains of a peer review. I am also very grateful to Marion who accepted to be a member of the jury although she is very busy with her second baby!

The faculty at Sciences Po was very helpful and I benefited from an excellent working environment. In particular, I would like to thank: Florian Oswald, who helped me cope with the struggles of structural labor; Eduardo Perez-Richet, who gave me the chance to teach Graduate Microeconomics and Market Design; Emeric Henry, who trained me and advised me before and during the job market. I would also like to thank the administrative staff, notably Alain Besoin, Pilar Calvo, Claudine Lamaze and Sandrine Le Goff.

Many others provided useful feedback at different stages of my PhD: among them there are the young labor economists of our PhD program, Nicolò Dalvit, Julien Pascal, Aseem Patel, Joanne Tan and Riccardo Zago. In these years, I also benefited from the advice of several experienced researchers, and I am particularly grateful to Arnaud Dupuy, Chris Flinn, Maurizio Mazzocco, Costas Meghir, Nicola Persico, Bernard Salanié and Frederic Vermeulen. In particular, I would like to thank Arnaud and Frederic for accepting to be members of the jury and for their useful comments on my manuscript.

I will soon join Frederic in Leuven with his colleagues Laurens Cherchye, André Decoster, Ferre De Graeve, Bram De Rock, Michal Kobielarz and Frank Verboven, who were all very supportive during the job market: I really look forward to this new experience.

After four years spent in our attractive open-space “les combles”, I am now ready to say goodbye to the younger recruits who will stay a little longer. I would like to thank all those who shared this experience with me and those who helped me during the months of the Job Market. A special thanks goes to my travel mates Julien, Max, Nicolò and Pippo. I am glad I could share many meals (some of dubious quality) and/or football games (virtual or real) with nice colleagues such as Aseem, Arthur, Assia, Charles, Clément, Elisa, Etienne, Ivan, Jan, JB, Jean-Louis, Joanne, Julia, Lucas, Ludovic, Nico, Oliver, Odran, Paul, Pauline, Pierre C, Pierre D, Riccardo, Sophie, Stefan, Tyler, Vladimir and Zydney. Good luck for the future!

It is also time for me to leave Paris and to say goodbye to all the people that I met or got to know better during these years. First of all, my numerous roommates: Nicolò, who holds a record of three mentions in these acknowledgments; Griso, who took me to the Bahamas; Barto, who did not formally live with us but was always very welcome; Louis, who watched so much Italian football with us; Rosi, who stayed in Paris longer than expected in her long wanderings; and all the other members of our cheerful community at Rue du Château d’Eau: Cassandra, Chiara, Fabio, Joschka, Ludovica and Mirza. Thanks also to all the friends that I have met in Paris: Ale, Alena, Antoine, Arianna, Baptiste, Bojana, Bruno, Dario, Edo Zanuttini, Emilien, Flavia, Fra Granella, Fra Rochetti, Hélène, Lisa, Martina, Reggio, Simon and many others.

When I first moved to Paris, I left behind many friends in Italy: since then, many have moved abroad, some have become investment bankers and engineers, others architects and rockstars. I am very proud of all of them. Thanks to all those who visited me, invited me over and welcomed me back, or simply made an effort to keep in touch during these years.

I am immensely grateful to Sibylle, who has been by my side since my very first days in Paris in spite of many little difficulties on our way. I am glad that we could work as a team, move in together and start planning our next move to Belgium. I would also like to thank her family and friends, who have always been very welcoming and ready to help.

Finally, I would like to thank my parents and my sister: the more time I spent away from home, the more I missed them, but their constant support from far away made my life so much easier in these years.

Introduction

Economists think of families as “complex decision units in which partners with potentially different objectives make consumption, work, and fertility decisions” ([Browning, Chiappori, and Weiss, 2014](#)). In many cultures, the institutions of marriage and divorce are the two most important means of acknowledging the formation and dissolution of the interpersonal ties that bind partners together. While the rights and obligations they establish change over time and space, marriage and divorce are both among the most defining events in one’s life and among the most complex economic decisions, with far-reaching implications for one’s working career, health and happiness.

Whether an individual is married, divorced, cohabiting or single matters for many choices and outcomes at the microeconomic level. The economic analysis of both labor supply and demand for consumption and investment goods requires a good understanding of how families work. On top of noneconomic benefits associated with companionship and intimacy, marriage and cohabitation are associated with the sharing of public goods, such as childcare, housing, related expenses and other durables. Partners can share risk by diversifying income sources, making joint saving decisions and adjusting labor supply ([Blundell, Pistaferri, and Saporta-Eksten, 2016](#)); they can also coordinate time schedules to share the burden of time-intensive tasks (e.g., chores) or, conversely, specializing in specific tasks (e.g., one spouse works full time while the other takes care of the children). Importantly, the household decision process affects the allocation of resources across family members: in some countries, consumption inequality within the family has decreased over time; in others, householders are still able to divert most resources to meet their personal benefits.³ In addition, when it comes to decisions that have implications for the future - such as investment in human capital and fertility - individuals are influenced not only by their current marital status but also by their expectations about their future marriage prospects ([Chiappori, Salanié, and Weiss \(2017\)](#)). Gender asymmetries in roles within marriage can be held responsible for a non-negligible share of the gender wage gap ([Blundell, Costa Dias, Meghir, and Shaw, 2016](#)).

³[Lise and Seitz \(2011\)](#) show that, in the U.K., consumption inequality between spouses has decreased in the period from 1968 to 2001, partly offsetting the rise in consumption inequality across households. [Dunbar, Lewbel, and Pendakur \(2013\)](#) show that, in Malawi, husbands still appropriate a large share of household resources, while many wives and children are likely to live below the poverty threshold.

At the aggregate level, systematic differences in marriage and divorce rates across population groups are an important factor of inequality between households (Greenwood, Guner, Kocharkov, and Santos, 2016). The United States are the leading example of this emerging divide across education and income strata: individuals at the bottom of the distribution are not only less likely to get married, but also more exposed to the risk of divorce and more likely to give birth out of wedlock (Carbone and Cahn, 2014); cohabitation is increasingly more common among people with low income and is associated with lower commitment with respect to legal marriage (Lundberg and Pollak, 2014). In addition, people exhibit a strong tendency to marry their likes: this typically results in the wealthy marrying the wealthy and the poor marrying the poor, with important implications for inequality between households (Fernández and Rogerson, 2001; Fernandez, Guner, and Knowles, 2005).

Ultimately, one of the main reasons why economists and policymakers are interested in marriage is that the family is the place where parents raise and invest in children. The characteristics and actions of parents have a tremendous impact on child development. We have learned that higher parental income is associated with lower borrowing constraints, better schools and a safer environment where to grow up (Acemoglu and Pischke, 2001); but parenting also has a direct influence on children's human capital endowment and helps shape cognitive and non-cognitive skills during childhood (Cunha, Heckman, and Schennach, 2010). Much of the variability in individuals' outcome, particularly lifetime earnings, is due to attributes that are determined during childhood (Heckman and Mosso, 2014). Differences in family background result in inequality of opportunities among children within the same cohort and are key determinants of intergenerational mobility (Becker and Tomes, 1979). Importantly, children raised by single parents are likely to lag behind their peers (McLanahan, 2004). Finally, the socioeconomic environment can, in its turn, influence the choice of parenting style, as in societies characterized by a higher degree of inequality parents have a stronger incentive to heavily invest in children in order to increase their odds of being successful (Doepke and Zilibotti, 2017).

The 1992 Nobel prize Gary Becker provides us with powerful analytical tools based on rational choice theory to study these issues and address the following fundamental questions (Becker, 1973, 1974, 1981): why do people get married, how do they choose their partner, and why do sometimes marriages end in divorce? A first key insight of Becker's approach is that, at least in many modern economies, marriage is a matching game: people are free to choose if and whom to marry, but only to the extent that they can successfully court and propose to the partner they desire. The nature of the gains from marriage can thus rationalize the aggregate marriage patterns that we observe in the data. For instance, when human capital is a valuable input for raising children, individuals compete for partners with a high level of schooling: as a result of this competition, the high educated get married with their likes, and so do the low educated. In reality, when

seeking a partner, individuals face obstacles of several kind, as search takes time and effort, while one's network of acquaintances is typically limited. However, in spite of these "frictions", Becker's key insight can be extended as long as individuals are left with some freedom to choose (Shimer and Smith, 2000).

A second key implication of Becker's theory is that one's bargaining power within the couple is proportional to how attractive she is on the marriage market: if an individual has many prospective spouses, she can afford being more demanding once engaged. Finally, Becker observes that divorce occurs when the gains from marriage fade: spouses might go through ups and downs and eventually realize that they are better off alone or with someone else (Becker, Landes, and Michael, 1977). However, unemployment and income shocks can also severely threaten the stability of the couple (Weiss and Willis, 1997): this is a major area of intervention for policymakers.

In this dissertation, I combine these key insights from Becker's theory with recent advances in the econometrics of matching models to address some of the new challenges that the economics of the family needs to face in the 21st century. The three essays aim to shed light on some key evolutions that the institution of marriage has experienced in the United States over the last 60 years. Same-sex unions have only been recently legalized in the United States and in several other Western countries, but still face important biological and legal constraints in childbearing: how do gains from marriage for same-sex couples differ with respect to different-sex couples? The American population is nowadays older, more educated and ethnically mixed than it used to be: are people also more open to the possibility of marrying someone who does not share the same ethnic or socioeconomic background? While wage inequality has increased since the 1980s, the gender wage gap has been shrinking: how has this changed the gains from marriage and the inequality patterns between households?

In the first chapter, we discuss the differences between same-sex and different-sex couples using a representative sample of Californian households for the period from 2008 to 2012. Our first contribution is to extend Becker's analytical tools to the study of same-sex marriage. Partners engaged in different-sex unions come from two separate populations (male and female) that sometimes present important asymmetries (e.g., the distribution of earnings and anthropometric characteristics). Conversely, we argue that matching among people of the same sex is a game that is known in matching theory as the "roommate problem": people are free to form pairs within the same group; namely, in this case, people are free to choose a partner of the same sex. While the identification and estimation of the gains from marriage for different-sex couples is a widely studied problem (Choo and Siow, 2006), we propose to extend these econometric tools - and in particular those developed by Dupuy and Galichon (2014) - to the analysis of the "roommate problem" in order to study same-sex marriage.

In our empirical analysis, we estimate the degree of complementarity and substitutability between observable traits for a sample of Californian same-sex couples from the 2008-2012 American Community Survey. In line with Becker’s theory, our estimates are both informative about the source of the gains from marriage and useful in measuring the strength of attraction between likes on the marriage market. The identification strategy relies on the assumption that we are able to observe the stable marriage market outcome in the data and thus infer the underlying marriage payoffs. Hence, the estimation consists in recovering the structure of payoffs that can best explain the marriage patterns in the data.

We find that different-sex couples are much more strongly segregated with respect to ethnic background with respect to both male and female gay couples. Aside from ethnicity, education is the most important dimension of sorting for both male and female gay couples; age also matters, but only comparatively less. In contrast, age is the most important dimension of sorting for different-sex couples, with education being the second most important. In addition, same-sex couples seem to be characterized by a lesser degree of household specialization: partners sort positively on hourly wages and hours worked; in comparison, partners in different-sex couples exhibit weak but positive sorting on wages, but negative sorting on hours worked. Finally, our model based on matching on observable demographic traits is better at explaining marriage patterns among different-sex couples than among same-sex couples. In other words, we find that unobservable traits are comparatively more important for same-sex couples.

Our analysis has the merit of providing a clear-cut comparison of sorting patterns across marriage markets that differ both for theoretical (two-sided matching vs “roommate problem”) and empirical reasons (the demographic composition of the two populations is very different; e.g., individuals engaged in same-sex couples are on average more educated). On the other hand, we do not attempt to delve further and explain what structural differences in behavior yield this heterogeneity in the structure of complementarities. Nevertheless, we run a series of robustness checks in order to show that differences in fertility might be able to explain much of the differences we observe in marriage patterns across these markets. We hope that, in the future, improvements in the quality of the data on same-sex marriage and advances in our understanding of fertility choices will allow us to address this kind of questions.

In the second chapter, we discuss how the nature of the gains from marriage has evolved in the United States from 1964 to 2017 and the implications of such changes for cross-sectional income inequality. In light of Becker’s theory, studying the evolution of marriage patterns is primarily insightful because it provides insights about changes in the institution of the family and the nature of the gains from marriage. For instance, ([Becker, 1973](#)) suggests that spouses sort negatively with respect to hourly wages if the gains from

marriage come from household specialization: in our empirical analysis, we can test whether this prediction holds in the 1970s as well as today. Moreover, the way people sort into couples on marriage markets is a determinant of economic inequality between households, particularly when wages and schooling are important sorting dimensions (Burtless, 1999; Fernandez, Guner, and Knowles, 2005): in our analysis, we test whether the strength of sorting with respect to these variables has increased over time.

Our first contribution is to draw from the literature on the estimation of matching model in order to disentangle between changes in the structure of the gains from marriage and compositional changes in the population. Differences in marriage patterns over time are the result of two main forces: the changes in the gains from marriage due to evolving household technology, gender roles, legal framework, etc.; and the changes in the demographic and socioeconomic composition of the population due to aging, increased ethnic diversity, increased college attendance, etc. For our analysis, we use the Current Population Survey data spanning from 1964 to 2017. We once again rely on the estimation technique proposed by Dupuy and Galichon (2014) and estimate the degree of complementarity and substitutability for a rich set of matching variables and for each yearly cross-section. We use our estimates of the complementarity parameters as measures of the strength of sorting and to discuss the changes in the nature of the gains from marriage.

We find that the importance of education as a sorting dimension has increased since the 1960s, while the importance of age has decreased. Racial segregation on marriage markets used to be extremely strong in the 1960s, but has much decreased in the 1970s and is nowadays slowly increasing: this recent trend seems to be driven by an increasing divide between Whites and Hispanics. While we find negative sorting with respect to hourly wages in the 1960s consistently with the predictions of Becker (1973), we observe an emerging complementarity between the spouses' earning potentials starting from the 1970s.⁴

In the second part of the paper, we run a counterfactual experiment where we hold the structure of complementarities in marriage - i.e., the “marital preferences” - fixed at its 1971 levels, while we let the demographic composition of the American population change as observed in the data. We find that, had marital preferences not changed, the 2017 Gini coefficient between married households would be lower by 6%. We conclude that about 25% of the increase in income inequality between married households is due to changes in marital preferences. We decompose this result and show that the increase in educational sorting plays a minor role in explaining the rise of inequality, while the emerging complementarity in hourly wages is one of the main driving forces.

⁴Depending on the measure of labor earnings that we use, we find weakly negative or weakly positive sorting with respect to this dimension. We use several measures of labor earnings to indirectly address the issue of endogenous labor market participation choices. For any measure, however, we observe an increasing degree of sorting on labor earnings.

In the third chapter, I study how changes in the wage distribution have reshaped marriage patterns in the United States between the 1970s and the 2000s and discuss their implications for the welfare of different population groups. Until 1970s, a vast majority of Americans got married at a young age and enjoyed stable relationships regardless of their socioeconomic status. Since then, the overall share of married adults has considerably declined, with the decrease being stronger for individuals without a college degree and with a low wage. These same individuals, once in a relationship, are also more exposed to the risk of divorce.

In order to understand these changing patterns and this emerging cleavage, I build an equilibrium search-and-matching model of marriage and divorce over the life-cycle in the wake of [Shimer and Smith \(2000\)](#) and [Chade and Ventura \(2002\)](#). In the model, singles look for a partner while facing search frictions. Upon a meeting, individuals decide whether to marry or not: the motives to get married are both economic and noneconomic. Spouses enjoy the company of a partner due to both affinity by age and education and idiosyncratic reasons (e.g., love and feelings): the latter are volatile, as couples go through ups and downs. Married agents can allocate a part of their time and earnings to the production of a public good, which is a source of economies of scale. The labor supply of women is endogenous: some families will rely on two earners, others will adopt a breadwinner-caretaker scheme. In addition, by pooling resources together, spouses also have the possibility to insure each other against wage shocks. However, in lack of full commitment, both wage shocks and love shocks can cause spouses to break up. After a divorce, agents are free to look for a new spouse, although their marriage prospects change as they get older.

In the empirical analysis, I use data moments taken from the Current Population Survey and the Panel Study of Income Dynamics and estimate the model for two separate periods, the 1970s and the 2000s. The estimated model aims to replicate three main features of the data: the cross-sectional marriage patterns - who gets married, and with whom; the longitudinal marriage patterns - the odds of getting married and divorced at different stages of the life-cycle conditional on one's characteristics; and the female labor supply patterns. Following [Goussé, Jacquemet, and Robin \(2017\)](#), I discuss the following identification puzzle: if search frictions are present, it is hard to tell whether people match with their likes because they enjoy their company or because these are the only people they meet. Using panel data on the length of marriage and singlehood spells, it is possible to disentangle between the two alternative explanations. The key intuition behind this identification strategy is that both marriage and divorce rates contain information on the underlying value of marriage: the model allows me to establish a theoretical link between the two and to exploit their joint variation in the data.

After estimating the model for the two periods, I run a series of counterfactual experiments to provide a quantitative assessment of the role played by changes in different factors

in explaining changing marriage patterns. I find that changes in the wage distribution alone can explain about a third of the decline in the share of married adults between the 1970s and the 2000s. They can also partly explain the emerging gap between married adults between male college graduates and non-college graduates. As divorce is costly, adjustments mostly occur on the entry side: agents accept fewer proposals and search longer before getting married. In the counterfactual experiment, the shrinking gender wage gap triggers the following mechanism. After the changes in the wage distribution, high-wage men are still relatively successful on the marriage market: they can form two-earner households with high-wage women or specialized households with low-wage women. Conversely, low-wage men struggle more than they used to: they are unlikely to marry women that earn more than them as they are not fit as caretakers, while low-wage women are now earning almost as much as they do. As their comparative advantage within the household fades, low-wage men grow less likely to get married and more exposed to divorce. This is exacerbated by the fact that women have stronger incentives to look for a wealthy partner, as they gained economic independence and there is now more distance between the top and the bottom of the male wage distribution.

Finally, I build measures of intertemporal welfare for individuals at the beginning of their adult lives. These measures take into account both agents' human wealth - defined as the expected stream of labor market earnings that they are able to generate while they stand on their own - and the expected gains from marriage over the life-cycle. This second term accounts for the expectations about the number, the timing and the duration of future marriages and about the characteristics of future partners. I find that, following changes in the wage distribution, young men suffer, on average, a 10% decrease in welfare: I decompose this finding and show that the decrease is mainly due to worsened labor market prospects (8.5%), while the remaining part is due to worsened marriage prospects (1.5%). Surprisingly, women do not experience any welfare gain: while the improvements in labor market conditions they experience result, on average, in a welfare increase, the latter is offset by worsening marriage prospects.

The model proposed in the last chapter stands, jointly with the recent work of ([Shephard, 2018](#)), as one of the very first examples of empirically-tractable equilibrium models of the marriage market over the life-cycle. With improvements to the specification and to the estimation technique, it can be used in the future as a policy lab to understand how people adjust their life-cycle marriage behavior when changes in a primitive parameter kick in or when a new policy is introduced. An important instance is the case of means-tested transfers: these are often conditional on recipients' marital status and their introduction is likely to influence marriage and divorce decisions.

Bibliography

- ACEMOGLU, D., AND J.-S. PISCHKE (2001): “Changes in the wage structure, family income, and children’s education,” *European Economic Review*, 45(4-6), 890–904.
- BECKER, G. S. (1973): “A Theory of Marriage: Part I,” *Journal of Political Economy*, 81(4), 813–846.
- (1974): “A Theory of Marriage: Part II,” *Journal of Political Economy*, 82(2), S11–S26.
- (1981): *A treatise on the family*. Harvard University Press.
- BECKER, G. S., E. M. LANDES, AND R. T. MICHAEL (1977): “An economic analysis of marital instability,” *Journal of Political Economy*, 85(6), 1141–1187.
- BECKER, G. S., AND N. TOMES (1979): “An equilibrium theory of the distribution of income and intergenerational mobility,” *Journal of Political Economy*, 87(6), 1153–1189.
- BLUNDELL, R., M. COSTA DIAS, C. MEGHIR, AND J. SHAW (2016): “Female labor supply, human capital, and welfare reform,” *Econometrica*, 84(5), 1705–1753.
- BLUNDELL, R., L. PISTAFERRI, AND I. SAPORTA-EKSTEN (2016): “Consumption inequality and family labor supply,” *American Economic Review*, 106(2), 387–435.
- BROWNING, M., P.-A. CHIAPPORI, AND Y. WEISS (2014): *Economics of the Family*. Cambridge University Press.
- BURTLESS, G. (1999): “Effects of Growing Wage Disparities and Changing Family Composition on the US Income Distribution,” *European Economic Review*, 43(4), 853–865.
- CARBONE, J., AND N. CAHN (2014): *Marriage markets: How inequality is remaking the American family*. Oxford University Press.
- CHADE, H., AND G. VENTURA (2002): “Taxes and marriage: a two-sided search analysis,” *International Economic Review*, 43(3), 955–985.
- CHIAPPORI, P.-A., B. SALANIÉ, AND Y. WEISS (2017): “Partner choice, investment in children, and the marital college premium,” *American Economic Review*, 107(8), 2109–67.
- CHOO, E., AND A. SIOW (2006): “Who marries whom and why,” *Journal of political Economy*, 114(1), 175–201.
- CUNHA, F., J. J. HECKMAN, AND S. M. SCHENNACH (2010): “Estimating the technology of cognitive and noncognitive Skill Formation,” *Econometrica*, 78(3), 883–931.
- DOEPKE, M., AND F. ZILIBOTTI (2017): “Parenting with style: altruism and paternalism in intergenerational preference transmission,” *Econometrica*, 85(5), 1331–1371.
- DUNBAR, G. R., A. LEWBEL, AND K. PENDAKUR (2013): “Children’s resources in collective households: identification, estimation, and an application to child poverty in Malawi,” *American Economic Review*, 103(1), 438–71.
- DUPUY, A., AND A. GALICHON (2014): “Personality Traits and the Marriage Market,” *Journal of Political Economy*, 122(6), 1271–1319.

- FERNANDEZ, R., N. GUNER, AND J. KNOWLES (2005): “Love and money: A theoretical and empirical analysis of household sorting and inequality,” *The Quarterly Journal of Economics*, 120(1), 273–344.
- FERNÁNDEZ, R., AND R. ROGERSON (2001): “Sorting and long-run inequality,” *The Quarterly Journal of Economics*, 116(4), 1305–1341.
- GOUSSÉ, M., N. JACQUEMET, AND J.-M. ROBIN (2017): “Marriage, labor supply, and home production,” *Econometrica*, 85(6), 1873–1919.
- GREENWOOD, J., N. GUNER, G. KOCHARKOV, AND C. SANTOS (2016): “Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation,” *American Economic Journal: Macroeconomics*, 8(1), 1–41.
- HECKMAN, J. J., AND S. MOSSO (2014): “The economics of human development and social mobility,” *Annual Review of Economics*, 6(1), 689–733.
- LISE, J., AND S. SEITZ (2011): “Consumption inequality and intra-household allocations,” *The Review of Economic Studies*, 78(1), 328–355.
- LUNDBERG, S., AND R. A. POLLAK (2014): “Cohabitation and the uneven retreat from marriage in the United States, 1950–2010,” in *Human capital in history: The American record*, pp. 241–272. University of Chicago Press.
- MCLANAHAN, S. (2004): “Diverging Destinies: How Children are Faring under the Second Demographic Transition,” *Demography*, 41(4), 607–627.
- SHEPHARD, A. (2018): “Marriage market dynamics, gender, and the age gap,” .
- SHIMER, R., AND L. SMITH (2000): “Assortative matching and search,” *Econometrica*, 68(2), 343–369.
- WEISS, Y., AND R. J. WILLIS (1997): “Match quality, new information, and marital dissolution,” *Journal of Labor Economics*, 15(1, Part 2), S293–S329.

CHAPTER 1

Like attract like? A structural comparison of homogamy across same-sex and different-sex households¹

Abstract

In this paper, we extend Gary Becker’s empirical analysis of the marriage market to same-sex couples. Becker’s theory rationalizes the well-known phenomenon of homogamy among different-sex couples: individuals mate with their likes because many characteristics, such as education, consumption behaviour, desire to nurture children, religion, etc., exhibit strong complementarities in the household production function. However, because of asymmetries in the distributions of male and female characteristics, men and women may need to marry “up” or “down” according to the relative shortage of their characteristics among the populations of men and women. Yet, among same-sex couples, this limitation does not exist as partners are drawn from the same population, and thus the theory of assortative mating would boldly predict that individuals will choose a partner with nearly identical characteristics. Empirical evidence suggests a very different picture: a robust stylized fact is that the correlation of the characteristics is in fact weaker among same-sex couples. In this paper, we build an equilibrium model of same-sex marriage market which allows for straightforward identification of the gains to marriage. We estimate the model with 2008-2012 ACS data on California and show that positive assortative mating is weaker for homosexuals than for heterosexuals with respect to age and race. Our results suggest that positive assortative mating with respect to education is stronger among lesbians, and not significantly different when comparing gay men and married different-sex couples. As regards labor market outcomes, such as hourly wages and working hours, we find some indications that the process of specialization within the household mainly applies to different-sex couples.

¹Chapter coauthored with Alfred Galichon and Marion Goussé.

1.1. Introduction

How individuals sort themselves into marriage has important implications for income distribution, labor supply, and inequality (Becker, 1973). Strong evidence shows that assortative mating in marriages accounts for a non-negligible part of income inequality across households (Eika, Mogstad, and Zafar, forthcoming).

Individuals tend to mate with their likes, a pattern called *homogamy*. However, because of asymmetries between the distributions of the characteristics in male and female populations, homogamy cannot be perfect among different-sex couples. In other words, heterosexuals cannot always find a “clone” of the opposite sex to match with. A large body of the literature has noticed that, up until recently, “men married down, women married up” due to the sex asymmetry in educational achievement that has only recently started to fade (Goldin, Katz, and Kuziemko, 2006). Gender asymmetries exist in other dimensions such as biological characteristics (windows of fertility², life expectancy, bio-metric characteristics), psychological traits, economic attributes (due to the gender wage gap), ethnic and racial characteristics (immigration is not symmetric across sexes, see Weiss, Yi, and Zhang, 2013) or demographic characteristics (some countries, such as China, have comparatively more imbalanced gender ratios).

Homogamy has been famously studied by Becker’s seminal analysis of the family. Becker (1973) expects most non-labor market traits, such as “intelligence, height, skin color, age, education, family background or religion”, to be complements. However, he also suggests that some attributes could be substitutes; in particular, Becker suggests that we should observe a negative correlation between some labor market traits such as wage rates because of household specialization.³ In order to provide a structural explanation of homogamy, Becker proposed a model of positive assortative mating (PAM) in which men and women are characterized by some socio-economic “ability” index. In this model, the marriage market clears so that men are matched with women that are as close as possible to them in terms of this index, which subsumes all the characteristics that matter on the marriage market. The (strong) prediction of Becker’s model is that the rank of the husband’s index in the men’s population is the same as the wife’s in the women’s population. However, this does not imply that the partners’ indices are identical: they would be so only if the distributions of the indices were the same for both men and women’s populations.

²Women’s fertility rapidly declines with age, whereas men’s fertility does not. Biologists and anthropologists argue that this dissymmetry could explain the well-documented preference of men for younger women (Hayes, 1995; Kenrick and Keefe, 1992). Low (2013) evaluates this young age premium for women and names it “reproductive capital,” as it gives them an advantage on the marriage market over older women.

³Chiappori, Oreffice, and Quintana-Domeque (2012) model a Becker-like marriage market with sorting on a unidimensional index. The estimation of such index reveals that high values in some attributes can compensate for poor values in others, thus showing that sorting is based on trade-offs between traits.

This analysis of the marriage market has attracted wide attention in the economic literature, in spite of its shortcomings. One shortcoming is that it originally refers to different-sex unions only. However, in a growing number of countries, same-sex couples have gained legal recognition, and the institutions of civil partnership and marriage no longer require that the partners must be of opposite sex. This official recognition is the result of several legal disputes and social activism by the gay and lesbian communities.⁴ The issue of whether to recognize same-sex unions has long been a topical subject in many countries, since it challenges the traditional model of the family. From both an economic and a legal point of view, the definition of what “family” means has relevant political implications as long as this term is present - and is generally central - in many modern constitutions and legal systems. Consequently, family units benefit from a special attention of policy-makers. Therefore, a discussion of the issues related to same-sex marriage - remarkably at policy level - requires a good understanding of the similarities and differences in the household dynamics among same-sex and different-sex couples. Besides, it is important to remember that the legal recognition of same-sex couples is only one of many transformations that the institution of the family has gone through in the last decades ([Stevenson and Wolfers, 2007](#); [Stevenson, 2008](#)). Finally, since more and more data on same-sex unions have been made available, the extension of the economic analysis of family to the gay and lesbian population can now be taken to data.

While it is natural to consider an extension of Becker’s model to same-sex households, it is worth noting that the previous considerations on asymmetries between men’s and women’s distributions only hold as long as each partner comes from a separate set according to his/her sex. On the same-sex marriage market, the two partners are drawn from the same population and the distributions of the characteristics are the same. Hence, the assortative mating theory pushed to its limits implies that, in this setting, partners should be exactly identical, i.e., each individual will choose to marry someone with identical characteristics.

In spite of such theoretical predictions, facts suggest a very different picture. Recent empirical results on the 1990 and 2000 American Census show that same-sex couples have less correlated attributes than different-sex ones, at least in terms of a variety of non-labor traits, including racial and ethnic background, age and education ([Jepsen and Jepsen, 2002](#); [Schwartz and Graf, 2009](#)). Studies on Norway, Sweden ([Andersson, Noack, Seierstad, and Weedon-Fekjær, 2006](#)) and Netherlands ([Verbakel and Kalmijn, 2014](#)) led to similar findings. In order to explain these systematic differences, the literature has suggested several possible reasons. A first consideration is that gay people might be forced to pick from a restricted pool because of their smaller numbers in the population, thus having a narrower choice when selecting their partner, resulting in a more diverse

⁴Public actions for gay rights acknowledgment are often considered to have started in 1969, in New York City. See [Eskridge Jr \(1993\)](#) and [Sullivan \(2009\)](#) for a detailed history and a full overview of the arguments in favor of and against same-sex marriage.

range of potential matches (Harry, 1984; Kurdek and Schmitt, 1987; Andersson, Noack, Seierstad, and Weedon-Fekjær, 2006; Schwartz and Graf, 2009; Verbakel and Kalmijn, 2014). Furthermore, gay men and lesbians have been found to be more likely to live in urban neighborhoods than heterosexuals, and since diversities in socio-economic traits are stronger in cities, this facilitates the crossing of racial and social boundaries (Black, Gates, Sanders, and Taylor, 2002; Rosenfeld and Kim, 2005; Black, Sanders, and Taylor, 2007). In light of these observations, one could argue that the same-sex marriage market is faced with stronger search frictions. Nevertheless, this might not necessarily be the case if the potential partners gather in specific locations, as it happens in cities and neighborhoods that are considered “gay-friendly” (Black, Sanders, and Taylor, 2007). Other analysts argue that gay people may have different preferences than heterosexuals, as they tend to be less conservative than straight individuals. Some explanations in this regard point out that, since homosexuality is still considered in some cultures as at odds with prevailing social norms, gay men and lesbians might grow less inclined to passively accept social conventions, and consequently they would end up choosing their partner with fewer concerns about his/her background traits (Blumstein, Schwartz, et al., 1983; Meier, Hull, and Ortyl, 2009; Schwartz and Graf, 2009). The detachment from the community of origin and the research for more tolerant surroundings have an influence both on values and social norms and on the heterogeneity of interpersonal ties.⁵

A part of these explanations has to do with individual preferences, whereas another part has to do with demographics, i.e., the distribution of the characteristics in the population. It is clear that the explanations listed in the former paragraph, while different in nature, are not mutually exclusive, but all contribute to a better understanding of the equilibrium patterns. For instance, a high correlation in education may arise from individual tastes (as individuals could find more desirable to match with a partner of similar educational background), but also from demographics (indeed, if some educational category represents a large share of individuals, this will increase odds of unions within this category, thus mechanically increasing the correlation in education). When comparing the heterosexual and homosexual population, this is particularly relevant as their compositions present significant differences (e.g., gay people are, on average, more educated than heterosexuals; see Black, Sanders, and Taylor, 2007).

In this paper, we focus on differences in marital gains: we would like to compare the structure of complementarity and substitutability across same-sex and different-sex households. In order to do so, we need a methodology that helps us interpret the observation of

⁵Note that household location choice and social norms are strictly related: it has been reported that gay people often leave their town of origin and escape social pressure exerted by relatives and acquaintances and go living in larger cities reputed to be gay/lesbian-friendly (Rosenfeld and Kim, 2005). Analogously, they are aware that they have more probabilities of avoiding discrimination by achieving higher educational levels and orienting their professional choices toward congenial working environments (Blumstein, Schwartz, et al., 1983; Verbakel and Kalmijn, 2014).

matching patterns and disentangle the role played by household interactions from external demographic factors. This is achieved through a structural approach, which allows us to estimate the parameters of the marital surplus function in order to fit the patterns actually observed. Hence, this approach will require an equilibrium model of matching.

In the wake of [Becker \(1973, 1991\)](#), the economic literature has modeled the marriage market as a bipartite matching game with transferable utility. A couple consists of two partners coming each from a separate or identical subpopulation (respectively, in the case of different-sex and same-sex unions). Both partners are characterized by vectors of attributes, such as education, wealth, age, physical attractiveness, etc. It is assumed that, when two partners with respective attributes x and y form a pair, they generate a surplus equal to $\Phi(x, y)$, which is shared endogenously between them. In the case of separate subpopulations (different-sex marriage), the landmark contribution of [Choo and Siow \(2006\)](#) showed that the surplus function Φ can easily be estimated based on matching patterns modulo a distributional assumption on unobservable variations in preferences, and was followed by a rich literature ([Fox, 2010](#); [Galichon and Salanié, 2014](#); [Chiappori, Salanié, and Weiss, 2017](#), to cite a few). [Dupuy and Galichon \(2014\)](#) extended Choo and Siow’s model to the case of continuous attributes and propose the convenient bilinear parameterization $\Phi(x, y) = x' Ay$, where A is a matrix called “affinity matrix” whose terms reflect the strength of assortativeness between two partners’ attributes. However, the bipartite assumption is restrictive and does not allow to estimate the surplus on same-sex marriage markets, and, to the best of our knowledge, no such estimation procedure is proposed in the literature. In a theoretical paper, [Chiappori, Galichon, and Salanié \(forthcoming\)](#) focus on stable matchings in a finite population and show that, when the population to be matched is doubled by cloning, the same-sex marriage problem, or “unipartite matching problem” can be mathematically reformulated as a heterosexual matching problem, or “bipartite matching problem”.⁶ In section 1.2 of the present paper, we apply an analogous reasoning to the empirically tractable, large-population, two-sided matching model of [Dupuy and Galichon \(2014\)](#), in order to adapt their empirical strategy to the same-sex marriage market.

A few papers already deal with the issue of assortativeness among same-sex households, although none of them allows to draw conclusions on the structural parameters of the surplus function that drives the assortativeness. The most relevant benchmarks for the empirical results of this work are the aforementioned [Jepsen and Jepsen \(2002\)](#) and [Schwartz and Graf \(2009\)](#). Both papers make use of the American census data (1990 and 1990-2000 respectively) and find that members of different-sex couples are more alike than those of same-sex ones with respect to non-labor market traits. The heterogeneity

⁶In another recent theoretical work, [Peski \(2017\)](#) extends the NTU framework of [Dagsvik \(2000\)](#) and [Menzel \(2015\)](#) and discusses the existence of stable matching in the unipartite case. [Fox \(2018\)](#) proposes an empirically tractable TU framework that generalizes both the bipartite and the unipartite case, and applies it to the car parts industry.

in assortativeness is measured in a logit framework containing dedicated parameters for homogamy. In general, in a logit framework individuals choose their best option among all possibilities. However, this fails to take into account the fact that matching takes place under scarcity constraint on the various characteristics. In the present paper, we estimate a model of matching in which agents compete for a partner; our measures of assortativeness are given by the parameters of the surplus function in each market (gay, lesbian and heterosexual).

The contributions of the present paper are twofold. On a methodological level, this paper is the first to propose a structural estimator of the matching surplus which applies to same-sex households, or, more generally, to instances of the unipartite matching problem. On an empirical level, we provide evidence by means of a structural analysis that, as concerns age and ethnicity, different-sex couples exhibit a higher degree of assortativeness than same-sex ones. While we find, in line with previous results, that sorting on education is stronger among lesbians with respect to different-sex couples, our results suggest that assortativeness on education is not significantly different when comparing gay male and married different-sex couples. Further, we also look at labor market traits such as hourly wages and working hours. Comparing assortativeness on labor market outcomes between same-sex and different-sex couples hints to different family dynamics and differences in the household specialization process. Finally, we briefly discuss the estimates of the mutually exclusive affinity indices obtained through our saliency analysis.

The rest of the paper is organized as follows. Section 1.2 will present the model and section 1.3 the estimation procedure. We describe our data in section 1.4 and our results in section 1.5. Section 1.6 concludes.

1.2. Model

In what follows, it is assumed that the full type of each individual, i.e., the complete set of all individual characteristics that matter for the marriage market (physical attributes, psychological traits, socio-economic variables, sex, sexual orientation, etc.), is fully observed by market participants. Each individual is characterized by a vector of observable characteristics $x \in \mathcal{X} = \mathbb{R}^K$, which constitutes his or her observable type. However, following [Choo and Siow \(2006\)](#), we allow for a certain degree of unobserved heterogeneity by assuming that agents experience variations in tastes that are not observable to the analyst, but are observable to the agents. In this paper, types are assumed to be continuous, as in [Dupuy and Galichon \(2014\)](#), hereafter DG, and [Menzel \(2015\)](#). Assume that the distribution of the characteristics x has a density function f with respect to the Lebesgue measure. Without loss of generality, the marginal distribution of the attributes is assumed to be centered, i.e. $\mathbb{E}[X] = 0$.

1.2.1. Populations. A *pair* is an ordered set of individuals, denoted $[x_1, x_2]$ where $x_1, x_2 \in \mathcal{X}$, in which the order of the partner matters, which implies that the pair $[x_1, x_2]$ will be distinguished from its inverse twin $[x_2, x_1]$. In empirical datasets, x_1 will often be denominated “head of the household” and x_2 “spouse of the head of the household” even though this denomination is used mainly for practical reasons and cannot be fully representative of the actual roles in the household.⁷ A *couple* is an unordered set of individuals (x_1, x_2) , so that the couple (x_1, x_2) coincides with the couple (x_2, x_1) . A *matching* is the density of probability $\pi(x_1, x_2)$ of drawing a couple (x_1, x_2) . Pairs $[x_1, x_2]$ and $[x_2, x_1]$ stand for the same couple, so that the density $\pi(x_1, x_2)$ is the sum of the density of $[x_1, x_2]$ and of the density of $[x_2, x_1]$, hence the symmetry condition $\pi(x_1, x_2) = \pi(x_2, x_1)$ holds. This symmetry constraint means that the position of the individual must not matter and thus that there are no predetermined “roles” within the couple that would be relevant for the analysis.⁸

We shall impose assumptions that will ensure that everyone is matched at equilibrium, hence the density of probability of type $x \in \mathcal{X}$ in the population is given by $\int_{\mathcal{X}} \pi(x, x') dx'$, which counts the number of individuals of type x matched either as the head of household in a couple $[x, x']$, or as the spouse of the head in a couple $[x', x]$. Thus, we are led to assume:

ASSUMPTION 1 (Populations). *The density $\pi(x, x')$ over couples satisfies $\pi \in \mathcal{M}^{sym}(f)$, where*

$$\mathcal{M}^{sym}(f) = \left\{ \pi \geq 0 : \left(\begin{array}{l} \int_{\mathcal{X}} \pi(x, x') dx' = f(x) \forall x \in \mathcal{X} \\ \pi(x_1, x_2) = \pi(x_2, x_1) \forall x_1, x_2 \in \mathcal{X} \end{array} \right) \right\}.$$

In contrast, in the classical bipartite problem, we try to match optimally two distinct populations (men and women) which are characterized by the same space of observable variables \mathcal{X} , and it is assumed that the distribution of the characteristics among the population of men has density f , while the density of the characteristics among the population of women is g . In this setting, the set of feasible matchings is typically given by:

$$\mathcal{M}(f, g) = \left\{ \pi \geq 0 : \left(\begin{array}{l} \int_{\mathcal{X}} \pi(x, y) dy = f(x) \forall x \in \mathcal{X} \\ \int_{\mathcal{X}} \pi(x, y) dx = g(y) \forall y \in \mathcal{X} \end{array} \right) \right\}$$

⁷We will come back in section 1.5.3 to this assumption that the roles of partners are exchangeable, which we test using a number of proxies for asymmetric household roles.

⁸Candelon and Dupuy (2015) extend Chiappori, Galichon, and Salanié (forthcoming)’s analysis to a model where agents form couples with endogenously assigned roles according to their characteristics. The model is applied to team formation in professional road cycling. Fox (2018) employs a very general many-to-many matching framework where agents self-select to be buyers or sellers upon a meeting. In both cases, hierarchy (leader vs assistants) or roles (buyers vs sellers) are clearly defined upon a match and observed in the data. This is unlikely to be the case when it comes to more complex and long-lasting relationships such as marriage.

Hence, $\pi \in \mathcal{M}^{sym}(f)$ if and only if $\pi \in \mathcal{M}(f, f)$ and $\pi(x_1, x_2) = \pi(x_2, x_1)$. Thus the feasibility set in the unipartite problem and in the bipartite problem differ only by the additional symmetry constraint in the unipartite problem.

1.2.2. Preferences. We now model preferences. Following DG, it is assumed that a given individual x does not have access to the whole population, but only to a set of acquaintances $\{z_k^x : k \in \mathbb{Z}_+\}$, randomly drawn, which is described below.

ASSUMPTION 2 (Preferences). *An individual of type x matched to an individual of type x' enjoys a surplus which is the sum of three terms:*

(i) *the systematic part of the pre-transfer matching surplus enjoyed by x from his/her match with x' , denoted $\alpha(x, x')$.*

(ii) *an endogenous utility transfer from x' to x , denoted $\tau(x, x')$. This quantity can be either positive or negative; we assume utility is fully transferable, hence feasibility imposes $\tau(x, x') + \tau(x', x) = 0$.*

(iii) *a “sympathy shock” $(\sigma/2)\varepsilon^x$, which is stochastic conditional on x and x' , and whose value is $-\infty$ if x is not acquainted with an individual x' . The quantity $\sigma/2$ is simply a scaling factor. More precisely, the set of acquaintances is an infinite countable random subset of \mathcal{X} ; it is such that (z_k^x, ε_k^x) are the points of a Poisson process on $\mathcal{X} \times \mathbb{R}$ of intensity $dz \times e^{-\varepsilon}d\varepsilon$.*

While the stochastic structure of the unobserved variation in preference described in part (iii) of Assumption 2 may appear complex, it is in fact a very natural extension of the logit framework to the continuous case, as we now argue. Indeed, it will imply that the individual maximization program of an agent of type x with this set of acquaintances is

$$\max_{k \in \mathbb{Z}_+} \alpha(x, z_k^x) + \tau(x, z_k^x) + \frac{\sigma}{2}\varepsilon_k^x, \quad (1.2.1)$$

where the utility of matching with acquaintance k yields a total surplus which is the sum of three terms, the systematic pre-transfer surplus, the transfer, and the sympathy shock. Define the systematic quantity of surplus at equilibrium U by

$$U(x, x') := \alpha(x, x') + \tau(x, x')$$

thus an individual of type x maximizes $U(x, z_k^x) + (\sigma/2)\varepsilon_k^x$ over the set of his/her acquaintances, which are indexed by k . This induces an aggregate demand over the type space. Indeed, it follows from the continuous logit theory initiated in [Dagsvik \(1994\)](#) that the conditional probability density of an individual of type x matching with a partner

of type x' is

$$\pi(x'|x) = \frac{\exp \frac{U(x,x')}{\sigma/2}}{\int_{\mathcal{X}} \exp \frac{U(x,x')}{\sigma/2} dx'}. \quad (1.2.2)$$

It is clear from expression (1.2.2) that this is a generalization of the logit framework to the continuous case.

Note that, by the property of independence of irrelevant alternatives (IIA) of the logit model, we do not need to describe the utilities of unmatched agents as long as the distributions of their stochastic parts are assumed to remain in the logit setting. Indeed, in the dataset we use, all agents are matched. Of course, one may worry about a potential equilibrium selection issue, i.e., that being matched affects the distributions of the agents' unobserved heterogeneity; however, in the logit setting, the IIA property guarantees that the distributions are preserved even after the selection, as shown in appendix D of DG. This is the reason why we consider a model where everyone is matched at equilibrium.

1.2.3. Equilibrium. Next, we define equilibrium in this framework. Denote

$$\Phi(x, x') := \alpha(x, x') + \alpha(x', x) = U(x, x') + U(x', x)$$

the systematic part of the joint surplus⁹ between x and x' . It follows from (1.2.2) and symmetry of π that

$$(\sigma/2) \ln \pi(x, x') = U(x, x') - a(x) = U(x', x) - a(x'), \quad (1.2.3)$$

$$\text{where } a(x) : = \frac{\sigma}{2} \log \int_{\mathcal{X}} \frac{1}{f(x)} \exp \frac{U(x, x')}{\sigma/2} dx'. \quad (1.2.4)$$

Substituting out for U in (1.2.3) yields the following equation, which expresses optimality in individual decisions:

$$\log \pi(x, x') = \frac{\Phi(x, x') - a(x) - a(x')}{\sigma}, \quad (1.2.5)$$

At equilibrium, the value of $a(\cdot)$ is determined by market-clearing condition $\int_{\mathcal{X}} \pi(x, x') dx' = f(x)$, that is

$$\int_{\mathcal{X}} \exp \left(\frac{\Phi(x, x') - a(x) - a(x')}{\sigma} \right) dx' = f(x). \quad (1.2.6)$$

We can now define our equilibrium matching concept.

DEFINITION 1. *The density π is an equilibrium matching if and only if there is a function $a(\cdot)$ such that both optimality equations (1.2.5) and market clearing equations (1.2.6) are satisfied.*

⁹Note that Φ is symmetric by definition, but α has no reason to be symmetric. Mathematically speaking, Φ is (twice) the symmetric part of α .

The main results on equilibrium characterization are summarized in the following statement:

THEOREM 1. *Under Assumptions (1) and (2):*

(i) *The equilibrium matching $\pi(x, x')$ is the unique solution to*

$$\max_{\pi \in \mathcal{M}(f, f)} \iint_{\mathcal{X} \times \mathcal{X}} \Phi(x, x') \pi(x, x') dx dx' - \sigma \mathcal{E}(\pi), \quad (1.2.7)$$

where $\mathcal{E}(\pi)$ is defined by

$$\mathcal{E}(\pi) = \iint_{\mathcal{X} \times \mathcal{X}} \pi(x, x') \ln \pi(x, x') dx dx'. \quad (1.2.8)$$

(ii) *The expression of $\pi(x, x')$ is given by*

$$\pi(x, x') = \exp\left(\frac{\Phi(x, x') - a(x) - a(x')}{\sigma}\right), \quad (1.2.9)$$

where $a(\cdot)$ is a fixed point of F , which is given by

$$F[a](x) = \sigma \log \int_{\mathcal{X}} \exp\left(\frac{\Phi(x, x') - a(x')}{\sigma}\right) dx' - \sigma \log f(x). \quad (1.2.10)$$

PROOF. By DG, Theorem 1, Problem (1.2.7) has a unique solution which can be expressed as

$$\pi(x, x') = \exp\left(\frac{\Phi(x, x') - a(x) - b(x')}{\sigma}\right)$$

for some $a(x)$ and $b(x')$ determined by $\pi \in \mathcal{M}(f, f)$. By the symmetry of Φ and by the symmetry of the constraints implied by $\pi \in \mathcal{M}(f, f)$, then $\tilde{\pi}(x', x) := \pi(x, x')$ is also solution to (1.2.7). By uniqueness, $\tilde{\pi} = \pi$, thus $\pi(x, x') = \pi(x', x)$. As a result, $b(x) = a(x)$ where a is determined by

$$\int \exp\left(\frac{\Phi(x, x') - a(x) - a(x')}{\sigma}\right) dx' = f(x)$$

QED. □

This result deserves a number of comments. First, we should note that there is an interesting interpretation of (1.2.7). While the first term inside the maximum tends to maximize the sum of the observable joint surplus, and hence draws the solution toward assortativeness, the second term $\mathcal{E}(\pi)$ is an entropic term which draws the solution toward randomness. The trade-off between assortativeness and randomness is expressed by the ratio Φ/σ . If this ratio is large, the assortative term predominates, and the solution will be close to the assortative solution. If this ratio is small, the entropic term predominates, and the solution will be close to the random solution. At the same time, note that the model parameterized by (Φ, σ) is scale-invariant: if $k > 0$, then the equilibrium matching

distribution π when the parameter is (Φ, σ) is unchanged when the parameter is $(k\Phi, k\sigma)$. This will have important consequences for identification, which is discussed in the next paragraph.

As a consequence of this result, we can deduce the equilibrium transfers and the utilities at equilibrium. Indeed, note that combining the expression of π as a function of U and a and equation (1.2.5) yields the following expression of U as a function of a :

$$U(x, x') = (\Phi(x, x') + a(x) - a(x')) / 2. \quad (1.2.11)$$

which is the systematic part of utility that an individual of type x obtains at equilibrium from a match with an individual of type x' . It is equal to half of the joint surplus, plus an adjustment $(a(x) - a(x'))/2$ which reflects the relative bargaining powers of x and x' . These bargaining powers depend on the relative scarcity of their types; indeed, $a(x)$ is to be interpreted as the Lagrange multiplier of the scarcity constraint which imposes that $\pi(\cdot, x)$ should sum to $f(x)$. Hence, the equilibrium transfer $\tau(x, x')$ from x to x' is given by

$$\tau(x, x') = (\alpha(x', x) - \alpha(x, x') + a(x) - a(x')) / 2. \quad (1.2.12)$$

Next, note that an interesting feature of Theorem A is that, while it characterizes equilibrium in the same-sex marriage problem, it highlights at the same time the equivalence with the different-sex marriage problem: indeed, as argued in DG, Theorem 1, the equilibrium matching in the different-sex marriage problem is given by the same expression as (1.2.7), with the only difference that $\mathcal{M}(f, f)$ is replaced by $\mathcal{M}(f, g)$, where f and g are respectively the distribution of men and women's characteristics.

We will use this characterization of the equilibrium matching as the solution of an optimization problem in order to estimate the joint surplus Φ based on the observation of the matching density π . As it is classical in the literature on the estimation of matching models with transferable utility, the primitive object of our investigations will be the joint surplus Φ rather than the individual pre-transfer surplus α ; indeed, without observations on the transfers, there is no hope to identify α : if we estimate that there is a high level of joint surplus in the (x, x') relationship, we will not be able to determine if this is due to the fact that “ x likes x' ” or “ x' likes x ”. We will only be able to estimate that there is a high affinity between x and x' .

1.3. Estimation

1.3.1. Estimation of the affinity matrix. Following DG, we assume a quadratic parametrization of the surplus function Φ to focus on a limited number of parameters

which could characterize the matching patterns. We parametrize Φ by an *affinity matrix* A so that

$$\Phi_A(x, y) = x' Ay = \sum_{ij} A_{ij} x^i y^j$$

where A has to be symmetric ($A_{ij} = A_{ji}$) in order for Φ to satisfy the symmetry requirement. Then the coefficients of the affinity matrix are given by $A_{ij} = \partial^2 \Phi(x, y) / \partial x^i \partial y^j$ at any value (x, y) . Matrix A has a straightforward interpretation: A_{ij} is the marginal increase (or decrease, according to the sign) in the joint surplus resulting from a one-unit increase in the attribute i for the first partner, in conjunction with a one-unit increase in the attribute j for the second. Hence, this approach is arguably the most straightforward way to model pairwise positive or negative complementarities for any pair of characteristics. It does, however, not preclude nonlinear functions of the x_i 's and the y_j 's, which can always be appended to x and y .

Recall equation (1.2.7), the optimal matching π maximizes the social gain

$$\mathcal{W}(A) = \max_{\pi \in \mathcal{M}(f, f)} \mathbb{E}_{\pi} [x' Ay] - \sigma \mathbb{E}_{\pi} [\ln \pi(x, y)] \quad (1.3.1)$$

which yields likelihood $\pi^A(x, x')$ of observation (x, x') , where π^A is the solution to (1.3.1). By the envelope theorem, $\partial \mathcal{W}(A) / \partial A_{ij} = \mathbb{E}_{\pi^A} [x^i y^j]$. Hence, our empirical strategy is to look for \hat{A} satisfying

$$\partial \mathcal{W}(\hat{A}) / \partial A_{ij} = \mathbb{E}_{\hat{\pi}} [x^i y^j], \quad (1.3.2)$$

where $\hat{\pi}$ is empirical distribution associated with the observed matching.

As noted before, the model with parameters (A, σ) is equivalent to the model with parameters $(kA, k\sigma)$ for $k > 0$. Hence, a choice of scale normalization should be imposed without loss of generality; a simple choice when a single market is considered is $\sigma = 1$, in which case the estimator A is meant as the estimator of the ratio of the affinity matrix over the scale parameter. The observation and comparison of multiple markets lead to slightly different normalization choices, which are discussed in section 1.3.5.

If a sample of size n , $\{(x_1, y_1), \dots, (x_n, y_n)\}$ is observed, then $\hat{\pi}(x, y)$ is the associated empirical distribution, which places mass $1/n$ to each observation. In DG, an estimator of A is obtained by solving the following concave optimization problem

$$\min_{A \in M_K} \mathcal{W}(A) - \mathbb{E}_{\hat{\pi}} \left[\sum_{ij} A_{ij} x^i y^j \right], \quad (1.3.3)$$

where M_K is the set of real $K \times K$ matrices. Indeed, the first order conditions associated to (1.3.3) are exactly given by (1.3.2). However, in the present case, the symmetry of A is a requirement of the model. The population cross-covariance matrix $\mathbb{E}_{\pi} [x^i y^j]$ is symmetric, as π satisfies the symmetry restriction $\pi(x, y) = \pi(y, x)$ in the population. Yet, in the

sample, $\hat{\pi}$ has no reason to be symmetric, as the first vector of variables x typically designates the surveyed individual, while the second vector of variables y designates the partner of the surveyed individual. Hence, the empirical matrix of covariances $\mathbb{E}_{\hat{\pi}}[x^i y^j]$ will only be approximately symmetric. Thus, we symmetrize the sample by adding the symmetric households, that is, if household ij is included, meaning that individual i was surveyed and reported partner j , we add a symmetric household ji , with j surveyed and reporting partner i . In other words, we replace the empirical distribution $\hat{\pi}(x, x')$ by its symmetric part $(\hat{\pi}(x, x') + \hat{\pi}(x', x))/2$. In the sequel, $\hat{\pi}$ will denote that symmetric part. This leads us to propose the following definition:

DEFINITION 2. *The estimator \hat{A} of the affinity matrix is obtained by*

$$\hat{A} = \arg \min_{A \in M_K} \{\mathcal{W}(A) - \mathbb{E}_{\hat{\pi}}[\sum_{1 \leq i, j \leq K} A_{ij} X^i Y^j]\}, \quad (1.3.4)$$

where M_K is the set of real $K \times K$ matrices.

The asymptotic behaviour of \hat{A} is computed in DG, theorem 2. A word of caution, is, however, in order. Although we have artificially doubled the sample size, by complementing household (x_i, y_i) with its mirror image (y_i, x_i) , one should beware that the sample size remains n , not $2n$. Thus, we can use directly the bipartite estimator on the mirrored sample, with the only modification that one will need to multiply the standard errors by a factor $\sqrt{2}$, as the effective sample size has not doubled.

1.3.2. Categorical variables. The previous analysis can be slightly adapted to deal with the case of categorical variables, such as race. Assume that the set of categories is denoted $\mathcal{R} = \{1, \dots, r\}$. Assume that the individuals are characterized by $x = (x^S, x^R)$, where $x^S \in \mathbb{R}^K$ are socio-economic characteristics, and $x^R \in \mathbb{R}^r$ is a vector of dummy variables x_i^R ($1 \leq i \leq r$) equal to 1 if individual x is of category $i \in \{1, \dots, r\}$, and zero otherwise. We work with the following specification of the surplus

$$\Phi(x, y) = (x^S)' A^S y^S + \lambda_R 1 \{x^R = y^R\} \quad (1.3.5)$$

where λ_R is a term that reflects assortativeness on the categorical variable, which provides a utility increment λ_R if both partners belong to the same category. Of course, this surplus function can be expressed multiplicatively as $\Phi(x, y) = x' A y$, where A can be written blockwise as

$$A = \begin{pmatrix} A^S & 0 \\ 0 & \lambda_R I_r \end{pmatrix} \quad (1.3.6)$$

and hence, A is obtained by running optimization problem (1.3.4) subject to constraint (1.3.6). Note that the envelope theorem implies that λ_R is identified by the moment matching condition

$$\Pr_{\pi}(x^R = y^R) = \Pr_{\hat{\pi}}(x^R = y^R)$$

which states that the predicted frequency of interracial couples should match the observed one.

1.3.3. Saliency analysis. The rank of the affinity matrix is informative about the dimensionality of the problem, that is, how many indices are needed to explain the sorting in this market. To answer this question, DG introduced *saliency analysis*, which consists of looking for successive approximations of the K -dimensional matching market by p -dimensional matching markets ($p \leq K$). Assume (without loss of generality as one can always rescale) that $\text{var}(X_i) = \text{var}(Y_j) = 1$. Then saliency analysis consists of a singular value decomposition of the affinity matrix $A = U'\Lambda V$, where U and V are orthogonal loading matrices, and Λ is diagonal with positive and decreasing coefficients on the diagonal. This idea is found in Heckman (2007), who interprets the assignment matrix as a sum of Cobb-Douglas technologies using a singular value decomposition in order to refine bounds on wages. This allows to introduce new indices $\tilde{x} = Ux$ and $\tilde{y} = Vy$ which are orthogonal transforms of the former, and such that the joint surplus reflects diagonal interactions of the new indices, i.e. $\Phi(x, y) = x'U'\Lambda Vy = \tilde{x}\Lambda\tilde{y}$.

Here, we need to slightly adapt this idea to take advantage of the symmetry of A and of the requirement that the matrix of loadings U and V should be identical. The natural solution is the eigenvalue decomposition of A , which leads to the existence of an orthogonal loading matrix U and a diagonal $\Lambda = \text{diag}(\lambda_i)$ with non-increasing (but not necessarily positive) coefficients on the diagonal such that

$$A = U'\Lambda U.$$

This allows us to introduce a new vector of indices $\tilde{x} = Ux$, which are orthogonal transforms of the previous indices. That way, the joint surplus between individuals x and y is given by

$$\Phi(x, y) = x'U'\Lambda Uy = \tilde{x}'\Lambda\tilde{y} = \sum_{p=1}^K \lambda_p \tilde{x}^p \tilde{y}^p$$

hence this term only reflects pairwise interactions of dimension p of \tilde{x} and \tilde{y} , which are either complements (if $\lambda_p > 0$) or substitute (if $\lambda_p < 0$), and there are no complementarities across different dimensions.

The following statement formalizes this finding:

THEOREM 2. *Assume that $\mathbb{E}_{\hat{\pi}}[X] = 0$ and that $\text{var}_{\hat{\pi}}(X^i) = 1$ for all i . Then there exists an orthogonal loading matrix \hat{U} and a diagonal $\hat{\Lambda} = \text{diag}(\lambda_i)$ with non-increasing coefficients on the diagonal such that*

$$\hat{A} = \hat{U}'\hat{\Lambda}\hat{U}$$

and, denoting $\tilde{x} = \hat{U}x$ and $\tilde{y} = \hat{U}y$, the estimator of the surplus function is given by

$$\hat{\Phi}(x, y) = \tilde{x}'\hat{\Lambda}\tilde{y} = \sum_{p=1}^K \lambda_p \tilde{x}^p \tilde{y}^p.$$

PROOF. Because \hat{A} is symmetric, it has the following eigenvalue decomposition

$$\hat{A} = \hat{U}'\hat{\Lambda}\hat{U}$$

where \hat{U} is orthogonal, and $\hat{\Lambda} = \text{diag}(\lambda_i)$ is diagonal with non-increasing coefficients. Denoting $\tilde{x} = \hat{U}x$ and $\tilde{y} = \hat{U}y$,

$$x' Ay = x' \hat{U}' \hat{\Lambda} \hat{U} y = \tilde{x}' \hat{\Lambda} \tilde{y} = \sum_{p=1}^K \lambda_p \tilde{x}^p \tilde{y}^p.$$

□

In the presence of categorical variables, the presence of a block λI_r in (1.3.6) reflecting assortativeness on the categorical variable implies that the singular values of A will be the singular values of \tilde{A} in addition to λ with multiplicity r . Therefore, it is recommended to perform saliency analysis simply on the upper left block \tilde{A} .

1.3.4. Selection issues. The purpose of the present paper is to compare match formation across same-sex and different-sex marriage market using the tools we developed above. In order to do so, we need to clearly delineate what is the relevant market in which agents match. We make the following assumption that gay men, lesbians and heterosexuals match on segmented markets, which we formalize into:

ASSUMPTION 3 (Exogenous selection). *The selection into either the same-sex or different-sex marriage market is exogenous.*

To relax this assumption, one would have to assume that all agents are pooled together in the same market, and choose their partner's gender among other characteristics, based on their own sexual orientation. The marital outcome, including the gender composition of households, would then be an equilibrium outcome resulting from a trade-off between socio-economic complementarities and other terms reflecting interactions between genders, and sexual orientations of the partners. We develop and discuss this relaxed framework in appendix 1.A, which shows that the market can be formulated as a single unipartite one where individuals are characterized by sexual orientation and gender in addition to other socioeconomic traits, and choose the gender of their partners among other characteristics. The more restrictive framework provided by assumption 3 can be obtained as a limiting case of the single-market framework where the interaction between sexual orientation and gender is predominant with respect to other characteristics, and thus the partner's gender is fully determined by own sexual orientation and gender.

In the absence of data on sexual orientation in our database, assumption 3 allows to infer sexual orientation from market participation, and therefore it permits to perform estimation of the affinity matrix expressing the interactions of the socioeconomic characteristics. However, if data on socioeconomic characteristics, gender *and* sexual orientation of matched partners were available, then the full matrix A could be estimated in a straightforward manner using our methodology, allowing to capture interactions not only between socioeconomic terms, but also between gender and sexual orientation, etc., as explained in appendix 1.A.¹⁰

As a final remark, the very fact that sexual orientation is exogenous is itself a strong assumption, and subject to current scientific debate. Researchers in biology, neuroscience, sexual medicine and psychology have provided evidence on the influence of psychological mechanisms on homosexual orientation (see [Jannini, Blanchard, Camperio-Ciani, and Bancroft, 2010](#); [Hines, 2011](#)). While there is an open debate among psychologists and social scientists about the stability of sexual behavior¹¹, research on early learning during childhood suggests that gender-typed behaviour –including sexual attraction– is internalized since infancy and stabilizes by late adolescence (see [Hines, 2011](#); [Dillon, Worthington, and Moradi, 2011](#)).

1.3.5. Comparison across markets. Affinity matrices are a useful tool to analyze marital surplus, and we would like to use them to compare sorting patterns across different-sex and female/male same-sex marriage markets. However, in order to achieve this, a discussion on normalization is needed. Indeed, recall from the above discussion that the equilibrium matching π is the solution to

$$\mathcal{W}(A, \sigma) = \max_{\pi \in \mathcal{M}(f, f)} \{ \mathbb{E}_{\pi} [X'AY] - \sigma \mathbb{E}_{\pi} [\ln \pi(X, Y)] \},$$

and therefore, a matching market with affinity matrix A and scaling parameter σ is observationally equivalent to another market with the same distribution of types and affinity matrix kA and scaling parameter $k\sigma$ for $k > 0$. Therefore, A and σ are not jointly identified, but only their ratio A/σ is identified.

It is therefore useful to adopt a normalization of (A, σ) . For cross-market comparison purposes, the normalization $\sigma = 1$ advocated in DG can be misleading, as it assumes that

¹⁰While this may be out of reach with current large-scale datasets, it is not unrealistic to believe that it will be possible to perform this type of analysis in the future. The National Survey of Family Growth, for instance, already contains detailed data on this topic, but unfortunately has no information on same-sex partnerships. This is due to the latter being recognized at federal level only recently.

¹¹[Diamond \(2008a\)](#) have provided the first piece of quantitative of the fluctuations of sexual orientation among adults women using long panel data. However, psychologists avoid talking about sexuality as a “choice of lifestyle”: [Diamond \(2008b, Chapter 5\)](#) considers changes in sexual orientation - and in other aspects of sexuality - as the consequence of “complex interplays among biological, environmental, psychological, and interpersonal factors”.

the standard deviation of the heterogeneity in preferences is the same across all markets considered. In this case, it seems more appropriate to normalize A by a factor so that the total quantity of surplus $\mathcal{W}(A, \sigma)$ is scaled to one in each market. That is:

ASSUMPTION 4. *The affinity matrix A and the amount of heterogeneity σ are normalized so that the equality $\mathcal{W}(A, \sigma) = 1$ holds in each market.*

When considering a single market, assumption 4 is a mere normalization, which can be imposed without loss of generality. It implies that the ratio of the average surplus provided by the interaction between characteristics i and j of two partners in a given market, divided by the average total surplus of a couple in that market is given by

$$\frac{A_{ij}\mathbb{E}[X^iY^j]}{\mathcal{W}(A, \sigma)} = A_{ij}\mathbb{E}[X^iY^j],$$

and hence $A_{ij}\mathbb{E}[X^iY^j]$ is the share of the average surplus explained by the interaction between characteristics i and j , relative to the average total surplus of a couple in that market.

On the contrary, when considering multiple markets, assumption 4 is no longer an innocuous normalization. It allows for a direct comparison of affinity matrices across markets, if one is willing to make the restrictive assumption that the average surplus of a couple is the same in every market. Note that, since $\mathcal{W}(kA, k) = k\mathcal{W}(A, 1)$ holds for any scaling parameter $k \geq 0$, in practice, we impose this normalization by first computing the estimator \hat{A} given by (1.3.4), and we then report $\hat{A}/\mathcal{W}(\hat{A}, 1)$.

1.4. Data

1.4.1. Data on same-sex couples. Empirical studies on same-sex marriage have traditionally needed to cope with poor data, due to the late legal recognition of these partnerships – still unachieved in several countries – and with misreporting issues, due to social pressure on respondents. Social scientists have largely relied on the data collected by the US Census Bureau for large sample analysis of same-sex unions (Jepsen and Jepsen, 2002; Black, Sanders, and Taylor, 2007; Schwartz and Graf, 2009). Starting from the 1990 decennial census, individuals could report themselves as “unmarried partners” within the household, regardless of their sex, so that gay couples could be identified. In more recent databases from the US Census Bureau, same-sex couples are still identifiable as out-of-marriage cohabiting partners. Indeed, although same-sex marriages have been officiated in some American states since 2004, they were recognized at federal level only in 2013, and currently available surveys conducted until then by the Census Bureau have not allowed reporting marriage bonds other than different-sex unions.

Accordingly, the present work relies on the five-year Public Use Microdata Sample (PUMS) for 2008-2012 coming from the ACS, conducted by the US Census Bureau. We restricted our sample to the state of California, which first legalized same-sex marriage on June 16, 2008 following a Supreme Court of California decision, and then – after some judicial and political controversies that impeded the officialization of same-sex weddings from November 5, 2008 to June 28, 2013.¹² – a decision of the U.S. Supreme Court finally accomplished full legalization. Restricting the sample to one state allows focusing on a marriage market with a uniform judicial framework. Moreover, in states where same-sex marriage is recognized, estimates on the number of married same-sex households are more reliable, and the incidence of the measurement error is smaller (Gates, 2010; Virgile, 2011)

1.4.2. Descriptive statistics. Our sample is limited to those individuals involved in a cohabiting partnership, both married and unmarried, thus excluding singles but also couples whose partners do not live in the same home. Each couple is identified as a householder with his/her partner, where both share the same ID household number.

The main database is composed of 681,060 individuals in couples who have completed their schooling. Because we restrict ourselves to prime age couples (both partners 25-50 year old), the size of our sample is decreased to 285,546 individuals. Out of them, 3,654 individuals (1.28% of the sample) live in same-sex couples, of which 2,034 male (0.71 %) and 1,620 female (0.57 %). 87.39% of the individuals in the sample are married heterosexuals and 11.33 % are cohabiting heterosexuals. For estimation purposes, after randomly selecting a subsample of different-sex couples¹³, a total of 9,820 couples are considered, of which 4,959 are married and 4,799 are not.

To compare different marriage markets, following Jepsen and Jepsen (2002), the main sample is divided into four subsamples: same-sex male couples, same-sex female couples, different-sex unmarried couples and different-sex married couples. This repartition is based on the assumption that individuals enter into separate markets according to their sexuality, in line with assumption 3. However, another criterion is used to differentiate two of the subgroups: married and unmarried different-sex couples are treated as two separate subpopulations¹⁴, since empirical evidence has reported significant differences in patterns between these two kinds of partnership (Jepsen and Jepsen, 2002; Schwartz and Graf, 2009). Although it is impossible to know *a priori* if a person is interested in a marital union rather than in a less binding relationship, this repartition can be of great interest and deepen the analysis. Nevertheless, even if California represents the larger state-level ACS sample in the US, further splitting the gay and lesbian groups into two subgroups would imply working with potentially very small samples. Moreover, although same-sex

¹²In this period, marriage licenses issued to same-sex couples held their validity.

¹³We randomly select 4% of married couples and 30% of unmarried couples.

¹⁴See Mourifié and Siow (2014) for a very interesting discussion of the endogenous choice of the form of marital relationship.

marriage is permitted, it has been recognized only recently and at the end of many legal struggles, which may have prevented a part of those same-sex couples that wished to marry from doing so. With more data available, considering married and unmarried same-sex couples separately would be extremely interesting, as proved by recent research of [Verbakel and Kalmijn \(2014\)](#) based on Dutch data.

Our study takes into consideration several variables, some related to the labor market and some others to the general background. Non-labor market traits include age, education and race. Age and education are treated as continuous variables, with the latter defined as the highest schooling level attained by the individual. Thanks to the detailed data of the ACS, the variable has been built in order to reflect as many distinct educational stages as possible. We consider five large racial/ethnic groups: Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Asian, Hispanic and Others.¹⁵ Finally, among labor market variables, we compute and include hourly wage¹⁶ and usual amount of hours worked per week. Note that yearly wage is top-coded for very high values (over \$999,999).

Table 1.1 presents some descriptive statistics of our sample. Individuals in same-sex couples are on average more educated than individuals in opposite-sex couples. As observed by [Black, Sanders, and Taylor \(2007\)](#), lesbians are much more likely to be part of the labor force than women in different-sex couples, and also have higher wages. We observe that unmarried different-sex couples are much younger than married couples and same-sex couples. Unmarried heterosexual men and women are on average four year younger than others. Cohabitation is often (but not always) a “trial” period before marriage, which can explain this age difference.¹⁷ Table 1.2 presents the distribution of ethnicity among couples: White individuals and Black women are overrepresented among lesbians, while Asians and Hispanics are under-represented in this population.

Table 1.3 presents correlations among traits. It shows that age and educational attainment are much more correlated among married different-sex couples than among unmarried and same-sex ones. Moreover, the correlation is stronger for lesbian couples than for gay male ones. Correlations on labor market outcomes are particularly interesting: hours worked are negatively correlated only for married different-sex couples, a possible clue of stronger

¹⁵American demographic institutions do not include a Hispanic category in variables on race, furnishing a separate variable for Hispanic origins, which is why there is some overlapping and the other categories bear the specification “Non-Hispanic”. The issue concerns the conceptual differences of “race” and “ethnicity”.

¹⁶The variable is computed as follows: we divide yearly wage by 52 in order to have the average weekly wage for last year and then we divide it again by the usual number of hours worked per week, which is available in the dataset. The hourly wage is partly approximated because the exact number of weeks worked in the last 12 months is not available. Note also that when information on labor earning or number of working weeks is missing, we set the hourly wage and the number of working weeks to 0 so that we keep all individuals in our analysis.

¹⁷This would require a dynamic framework, which we don’t have in our static model. See the theoretical work of [Brien, Lillard, and Stern \(2006\)](#) and [Gemici and Laufer \(2011\)](#).

household specialization, whereas the correlation is positive albeit low for same-sex couples. On the other hand, wages display a positive correlation in every market, with different-sex married couples and male same-sex couples exhibiting the lowest correlation.

Table 1.4a, 1.4b and 1.4c present homogamy rates of couples with respect to race for different types of couples. The homogamy rate is the ratio between the observed number of couples of a certain type and the counterfactual number which should be observed if individuals formed couples randomly. For instance, table 1.4b shows that lesbian couples among Black women form 10 times much more than if they were formed randomly among the lesbian population.

1.5. Results

Homogamy rates and correlations presented in section 1.4 are interesting measures of assortative mating and provide a good starting point for our analysis. However, they are not sufficient to reach any conclusion about the degree of assortativeness in the marriage market. By estimating the parameters of the surplus function, we compare the level of complementarity and substitutability between characteristics across different marriage markets. This approach is consistent with Becker's model of assortative mating, and allows us to measure the degree of assortativeness for each combination of characteristics *ceteris paribus*. In particular, we can test whether assortativeness on observables - notably, age, race and education - is weaker among same-sex couples, as found by [Schwartz and Graf \(2009\)](#).

While we measure the direction and strength of interactions between traits, we do not attempt to estimate preference and production terms separately. Hence, we cannot tell whether marital gains differ across markets because of differences in household production rather than pure taste for homogamy. In particular, we cannot tell to which extent differences in the opportunity cost of bearing and raising children affect sorting patterns.¹⁸ If couples wish to have genetically related children, Assisted Reproductive Technologies imply that children inherit genetic traits from only one out of two partners, with possibly important implications for sorting. In our empirical analysis, we limit ourselves to the estimation of the model on carefully chosen subsamples (e.g., childless couples) in order to provide an intuition of where the major sources of diversity between gay and different-sex couples lie. While not exhaustive, these robustness checks could constitute a useful starting point for future research.

¹⁸[Allen and Lu \(2017\)](#) propose a theoretical model which explains differences in expected matching behavior, marriage rates, non-child-friendly activities, and fertility, based on different costs of procreation and complementarities between marriage and children.

Finally, Becker’s model suggests that we interpret differences in assortativeness as a consequence of differences in marital gains, rather than as a consequence of search dynamics (notably, geographic factors and search frictions), segmentation into local markets along socio-economic traits, or preferences of third parties and social pressure (Kalmijn, 1998). In particular, gay individuals tend to move away from their hometowns and may not be “out” at school or in the workplace (Rosenfeld and Kim, 2005), and this could influence the composition of their interpersonal ties.¹⁹ Also in this case, some of our robustness checks can help understand how these concurrent forces affect our results. Nonetheless, we believe that a model explicitly accounting for such factors would be necessary to quantify their impact on sorting patterns.

We report in table 1.5 the estimates of the affinity matrix for gays, lesbians, married and cohabiting heterosexuals.

1.5.1. Age, education and race/ethnicity. Our estimates of the diagonal elements of the affinity matrices are highly positive and significant for age, education and ethnicity, which confirms previous findings about positive assortative mating. In line with the results by Jepsen and Jepsen (2002) and Schwartz and Graf (2009), we find that assortativeness on age and ethnicity is comparatively weak for male same-sex couples (0.62 for age, 0.62 for ethnicity), and progressively stronger for female same-sex (0.79, 1.26) and unmarried different-sex couples (1.14, 1.98), whereas married different-sex couples exhibit the strongest complementarities (2.17, 2.49). Results on education are more nuanced: complementarity of schooling levels is the strongest for lesbian couples (1.19), while estimates for married same-sex (0.82) and gay couples (0.84) are not significantly different. Finally, complementarity of schooling levels is the lowest for unmarried heterosexuals (0.66).

Our estimates on the level of educational sorting is partly at odds with previous findings. Empirical research on this topic mainly concluded that assortativeness on education is weaker on both male and female same-sex marriage markets with respect to different-sex marriage markets. However, the social science literature²⁰ provides a large set of explanations about why sorting patterns should differ across different-sex and same-sex couples, and not all of them predict that educational sorting is weaker among the latter. On the one hand, gay men and lesbians are expected to be more inclined to “transgress” social norms and to cross socio-economic and racial barriers when choosing their partner

¹⁹For instance, online dating among heterosexuals has been found to reduce assortative matching on education (Hitsch, Hortacsu, and Ariely, 2010). However, as dating apps and websites grow in popularity, thus giving access to a larger and larger pool of possible matches, and tend to specialize on segmented markets (e.g., by ethnicity or religion), we wonder if this conclusion still holds.

²⁰The main reference works about mating among gay people are listed in our introduction. We refer to Schwartz and Graf (2009) and Verbakel and Kalmijn (2014) who, drawing from literatures from different social sciences, both provide a complete and updated review on this topic.

(Rosenfeld and Kim, 2005; Schwartz and Graf, 2009). Our findings suggest that this effect might be prevalent as concerns age and ethnicity. On the other hand, gay people are also expected to have stronger *egalitarian preferences*. Verbakel and Kalmijn (2014) suggest that similar schooling levels can lead to a more equal division of labor. Spouses that aim to concentrate their efforts on the labor market rather than to specialize each in a different set of skills might thus exhibit a stronger level of assortativeness on education.

As anticipated above, childrearing is a major driver of household specialization, and same-sex couples are less likely to have children.²¹ Hence, we estimate the affinity matrix using the subsample of childless couples for each of our four marriage markets²² (see summary table 1.8 and the full tables in the online appendix). We find that, with respect to ethnicity, both childless same-sex and childless different-sex couples exhibit a weaker taste for homogamy compared with couples with children of the same respective sexual orientation. Similar results hold for sorting on age, although only differences between married different-sex couples with and without children are significantly different. It thus seems that individuals who plan to have children look for a more similar partner along these two dimensions than those who do not.

When it comes to education, the picture is a bit more contrasted. As for the previously discussed traits, one observes stronger assortativeness on education for same-sex couples with children than for those without. In contrast, childless different-sex couples are more assortatively matched on education than those with children. Different-sex couples who do not plan to have children will not benefit from large gains from specialization and may look for a partner with similar schooling. It is interesting to note that married childless different-sex couples are found to exhibit a higher degree of assortativeness with respect to age, ethnicity *and* education with respect to same-sex childless couples.²³

1.5.2. Labor market traits. To describe labor market traits, we must be very cautious as these outcomes are potentially endogenous. Since we do not observe these traits at the moment of the match formation but possibly much later, the specialization process at work in couples may have already begun. In particular, we expect that this specialization effect is strong in different-sex couples, who are more likely to have children.²⁴ Raising children takes time and many mothers leave the labor force or reduce

²¹In our sample, among the 25-50 years-old, 14.5% of gay men have children, 37.8% of lesbians, 58.04% of cohabiting different-sex couples and 83.5% of married different sex couples.

²²We are aware that the subsample constitutes an “artificial” marriage market, since individuals do not rigidly self-select into a separate market based on their preference for having children. However, our model does not have a specification that explicitly accounts for choices related to childbearing.

²³We also estimate the affinity matrix for married different-sex couples with one and three children (see table 1.8c). Our findings are in line with what stated in the main text. The higher the number of children, the stronger the assortativeness on age and ethnicity, and the weaker the assortativeness on education.

²⁴Antecol and Steinberger (2013) and Jepsen and Jepsen (2015) showed that to a lesser extent some household specialization also exists within same-sex households. Moreover, Antecol and Steinberger (2013) stress that childless different-sex couples are less specialized and thus more similar to same-sex couples.

their working hours. Consequently, because of interrupted careers and less paid part-time jobs, their hourly wage does not rise as much as that of their male counterparts and we observe many associations between low-wage women and high-wage men. This phenomenon could bias our estimates. To assess the importance and the sign of this bias, we also perform the estimation on four additional selected samples where the specialization effect should be limited : 1) childless couples, 2) bi-earner couples, 3) young couples (25-35 year old), 4) recently married couples with no children. The last selection is only available for different-sex married couples as we observe their wedding date; we keep couples who got married in the preceding year and who have no children. A summary of the results is available in tables 1.8a, 1.8b, 1.8c and 1.8d. The full tables are available in the online appendix. We first describe the general results obtained from the main sample.

First, we measure significant positive assortativeness on hourly wages for all types of couples, although the coefficient is higher for same-sex couples (0.05 for gays and 0.06 for lesbians) and for different-sex unmarried couples (0.05) than for married different-sex couples (0.01). Furthermore, we observe negative assortative mating on working hours for married different-sex couples (-0.04), whereas we observe much higher and significant positive estimates for same-sex couples (0.12 for gays and 0.20 for lesbians). The coefficient for unmarried couples is also positive and significant (0.09). Assortative mating on wages and working hours is likely to be related to the presence of children. As same-sex couples are less likely to have children, they have weaker incentives for specialization. Unmarried couples may also have lower preferences for children than married couples. To better understand this result, we estimate our model on childless couples. We find that married different-sex couples without children have a positive coefficient for both wage (0.07) and working hours (0.12), and thus are more similar to same-sex and different-sex unmarried couples. In this regard, our results are in line with those of [Jepsen and Jepsen \(2015\)](#). Similarly, the assortative mating coefficients for wages and hours are higher – but to a lesser extent – for same-sex couples without children compared to those with children. We also perform the estimation on married couples with only one child and married couples with three children to disentangle the effect of the presence of children among married couples. As couples have more children, we observe a decrease in assortative mating coefficients for wages and hours. Similarly, as expected, the estimation on bi-earner couples shows an increase in assortative mating on wages and hours.²⁵ In the case of young same-sex couples the comparison leads to less clear-cut conclusions due to small sample size, while positive assortative mating on labor market traits is stronger when comparing young different-sex couples with those in our main sample.

²⁵In our analysis, bi-earner couples are couples whose both members declare positive wage and number of working weeks.

The cross-estimate between the wage of one partner and working hours of the other partner is also very interesting to analyze. We find negative assortative mating on wages and working hours. The estimate is highly negative for lesbians (-0.19) and different-sex married couples (-0.09 for the interaction between wife’s wage and husband’s hours, -0.13 for the symmetric interaction). It is also negative and significant for gays (-0.07). Hence, for heterosexual married couples and homosexual couples, the match gain increases when one partner increases his/her wage and the other decreases his/her working hours. This result is robust to the presence of children, to the age of couples and to the bi-earner sample.

1.5.3. Cross-interactions and symmetry. Other significant positive cross-effects have been found for some off-diagonal elements of the affinity matrix. The parameter capturing the interaction between wage and education is persistently high and positive. This might suggest that higher wage individuals have a preference for more educated partners, keeping constant their wage and all other characteristics. Affinity between these two variables is relatively weaker for gay men (0.13) and lesbians (0.20). Complementarity between the two inputs is stronger for different-sex couples, but the relationship is asymmetric: estimates for married couples suggest that complementarity in husband’s wage and wife’s education is stronger (0.27) than the other way around (0.19). However, the corresponding estimates for unmarried couples go in the opposite direction (0.21 and 0.37 respectively). The complementarity between the two traits might be explained by the fact that high-income individuals - independently of their educational level - may enjoy the company of cultured partners.

Another cross-interaction that arises from the estimation is the substitutability between age and hours worked on same-sex marriage markets (-0.13 for gay men and -0.10 for lesbians). This interaction might be due to household bargaining dynamics, as explained by [Oreffice \(2011\)](#): younger partners enjoy higher bargaining power and thus can afford reducing their labor supply. Interestingly, unmarried different-sex couples exhibit similar patterns, although the effect is weaker.

Finally, as a last robustness check for our main results, we estimate a bipartite matching model of the same-sex marriage market where the affinity matrix is not required to be symmetric. In this case, we need to define two separate subpopulations to run a bipartite estimation, and therefore we need to define household roles. On one side of the market, we group all those partners that are registered as “householders”, whereas on the other we group their “cohabiting partners”. This repartition is highly artificial, since it implies that two gay individuals that are householders before finding a partner can never match: in general, it seems implausible to divide the same-sex population in two separate subgroups with the data that we have at hand. Nonetheless, it is interesting to check - under the strict assumption of predetermined roles - whether some asymmetry

in cross-interactions occurs. We observe that the affinity matrices for both gay men and lesbians (respectively tables 1.7a and 1.7b) are not much different than in the unipartite case. When testing for differences between the off-diagonal coefficients²⁶, we find that only the cross-interaction between age and hours is significantly different between householders and their partner in male same-sex households. The interaction is significant only in one direction, relatively young cohabiting partners can reduce their labor supply when matching with older householders. The difference is non-significant for female same-sex households. We already stated that we were not very confident in the label “householder” to define a particular role in the couple. As additional robustness checks, we assign a particular role to each partner in same-sex couples according to another characteristic. We test if there is a particular role assigned to 1) the older partner, 2) the higher earner (total income) partner. Each partner belongs to a certain population depending to his/her status in the couple, then we estimate a bipartite matching model of the same-sex marriage market on these two populations. Results are presented on the online appendix. Again, they are not much different from the unipartite case. We now detail some exceptions. When separate populations are defined according to the relative age, we now find a negative interaction between age of the older partner and education and hours worked of the younger partner among gay couples. When separate populations are defined according to the relative total income, symmetry is respected for all coefficients for male same-sex couples but not for female same-sex couples. Specialized roles may appear among lesbian couples as they are more likely to have children. There are asymmetries in the interaction between hours and education, hours and wage and education and wage. The coefficient of interaction between education and wage is always high and positive but it is much higher between the education of the higher earner and wage of the lower earner than the other way around.

1.5.4. Matching on unobservables. Thanks to assumption 4, we can evaluate the parameter σ for each market. As anticipated in section 1.2, this parameter has a simple interpretation: the higher σ , the more matching appears as random to the econometrician, or, in other words, the higher the entropy. Since our observable characteristics are meant to capture the main socio-economic traits, we expect that a higher σ implies that matching is less “deterministic”: indeed, for higher σ , the socio-economic background of an individual matters relatively less, whereas other unobservable traits (e.g. personality or physical appearance) may matter relatively more.

We find that entropy is higher on same-sex marriage markets (1.26 for gay men and 1.23 for lesbians), whereas it is lower on different-sex marriage markets (1.04 for unmarried couples and 1.00 for married). Hence, if we interpret entropy as due to the relevance of

²⁶We provide test results in the online appendix.

unobservables, it seems that, among same-sex couples, socio-economic background matters less relatively to unobservable traits.

While we privilege an interpretation of σ that is consistent with Becker’s frictionless matching model, it is important to recall that differences in the search process are also captured by σ . All else held constant, stronger search frictions should result in higher entropy (Shimer and Smith, 2000). However, we are not able to disentangle the effects of search frictions from the relevance of unobservables. This may be problematic when comparing the same- and different-sex marriage markets because gays constitute a relatively small part of the population. If the frequency of meetings is increasing in the size of the pool of potential partners, as suggested by the search-and-matching literature²⁷, then we should expect search frictions to be stronger for the gay population. In order to address this issue, we estimate the affinity matrix for a subsample of couples living in the metropolitan areas of Los Angeles and San Francisco. We expect search frictions to be small in densely populated urban areas, and meeting opportunities to be comparable for homosexuals and heterosexuals. While the point-estimates of σ are almost unchanged for different-sex couples (1.03 for unmarried, 1.00 for married), we observe an increase for both gay men (1.35) and lesbians (1.32).²⁸ Since the estimates of σ for same-sex couples are persistently higher even in areas where search frictions are expected to be lower, we conjecture that the difference in entropy can, at least in part, be explained by unobservable traits.

1.5.5. Saliency analysis. A way to bring further insights on the main drivers of preferences of individuals over different characteristics is to decompose the affinity matrix in orthogonal dimensions. As detailed in section 1.3.3, we conduct the decomposition analysis on all variables, with a specific treatment for race, a categorical variable. Under the parametrization (1.3.5), we estimate the affinity matrix driving the interactions of the non-race characteristics, and the coefficient λ_R which measures the homogamy on race. We obtain

$$\Phi(x, y) = \sum_{p=1}^K \lambda_p \tilde{x}^p \tilde{y}^p + \lambda_R 1 \{x^R = y^R\}$$

where \tilde{x}^p is the p -th index associated to the individual with characteristics x obtained by the decomposition, and x^R is the race characteristics. Taking expectations with respect

²⁷In the search-and-matching literature, the meeting rate is usually modeled as a constant return to scale function of vacancies of both sides of the market, usually a Cobb-Douglas function. This assumption seems to be supported empirically on the labor market (Petrongolo and Pissarides, 2001). On the marriage market, Goussé, Jacquemet, and Robin (2017) also use a Cobb-Douglas matching function.

²⁸Complete tables of results obtained with the subsample of couples residing in urban areas are available in the online appendix. Urban areas are defined as the Metropolitan Statistical Areas of Los Angeles and San Francisco.

to the sample distribution

$$\mathbb{E}_{\hat{\pi}} [\Phi (X, Y)] = \sum_{p=1}^K \lambda_p \mathbb{E}_{\hat{\pi}} [\tilde{X}^p \tilde{Y}^p] + \lambda_R \mathbb{E}_{\hat{\pi}} [1 \{X^R = Y^R\}],$$

this allows us to decompose the average surplus $\mathbb{E}_{\hat{\pi}} [\Phi (X, Y)]$ into the sum of the surplus created by the interaction between characteristics p , namely $\lambda_p \mathbb{E}_{\hat{\pi}} [\tilde{X}^p \tilde{Y}^p]$, plus the surplus created by homogamy on the racial characteristics $\lambda_R \mathbb{E}_{\hat{\pi}} [1 \{X^R = Y^R\}]$. We present such a decomposition in the appendix in table 1.6. First, we show that the share of the average surplus created by homogamy on the racial characteristics reaches more than 47% for different-sex couples, 42% for lesbian couples and only 33% for gays. Then, the rest of the average surplus can mainly be explained in two orthogonal dimensions which measure relative attractiveness. These indices load on different characteristics of individuals. For the different-sex and the lesbian marriage market, the first index is almost only composed of age. It explains by itself around 36% of the surplus for married different-sex couples, 33% for unmarried different-sex couples and 29% of the surplus for female same-sex couples. Then the second index mostly relies on education for these markets, and explains around 27% of the sorting for female same-sex couples, 20% for unmarried different-sex couples and 16% for different-sex married couples. For male same-sex couples, the first index of sorting relies on education (35%) then comes ethnicity, then age (30%). When we consider different-sex couples, the indices of mutual attractiveness could differ between genders. For married heterosexuals, the education/wage index (the second index) loads positively twice as much on wages for men than for women, whereas there is a penalty on working hours for women that does not appear for men.

1.6. Discussion and perspectives

The contributions of the present paper are twofold. From a methodological point of view, this paper is the first to propose a tractable empirical equilibrium framework for the analysis of same-sex marriage. Our methodology could be applied to many other markets (e.g. roommates, teammates, co-workers). In addition, we apply the model in order to provide an empirical analysis of sorting patterns in the same-sex marriage market in California. We conduct a cross-market comparison: we analyze the heterogeneity in preferences between same-sex and different-sex couples. First, we find that, as concerns age and ethnicity, the different-sex marriage market is characterized by a stronger “preference for homogamy” than the same-sex marriage market. Meanwhile, results are more nuanced when it comes to education: while lesbians show stronger assortativeness on education, there is no significant difference in that dimension between gay male and married different-sex couples. Second, we discuss the differences in complementarity and substitutability in the marriage surplus function as defined in Becker’s theory of the family. Our findings suggest that labor market traits are complementary only for same-sex and unmarried

cohabiting different-sex couples. The presence of children seems to be a central driver of these contrasted findings. These results indicate that the traditional concept of marriage gain based on specialization within the couple is still relevant today, although it applies mainly to couples with children.

Families and household arrangements are evolving quickly and we need to understand the underlying forces of these changes. The need for effective analytical frameworks to study and describe new forms of families has recently emerged in the economic literature, mainly as concerns same-sex couples (Black, Sanders, and Taylor, 2007; Oreffice, 2011) and cohabiting partners (Stevenson and Wolfers, 2007; Gemici and Laufer, 2011). Let us briefly discuss both. In this paper, we found structural differences between three separate subpopulations divided according to sexual preferences. However, can we state with certainty that these markets are mutually exclusive? In fact, individuals may endogenously choose into which market they are willing to match. The appendix explores a theoretical framework to move beyond the exogenous selection hypothesis. While the lack of individual-level data on sexual orientation does not allow at this stage to go beyond a theoretical model, we trust that, with the consolidation of the same-sex marriage and the availability of more and more accurate data, it will soon be possible to expand our understanding on these questions. Second, cohabitation is another developing phenomenon (Schwartz and Graf, 2009; Gemici and Laufer, 2011; Verbakel and Kalmijn, 2014) and is associated with a lower degree of specialization and a lower degree of positive assortative mating. A promising area of research would be to understand the preferences for marriage or cohabitation jointly with sorting preferences. Mourifié and Siow (2014) set a first model in that direction for different-sex couples, which could be adapted to same-sex couples using the techniques put forward in the present paper; see also the empirical analysis of Verbakel and Kalmijn (2014), and Aldén, Edlund, Hammarstedt, and Mueller-Smith (2015) for the marital and fertility decisions of same-sex households.

Another topic of interest is the effect of new forms of families on the traditional ones over time. Opponents of same-sex marriage have voiced the fear that it will cause the marriage institution to lose its value and favor alternative forms of families, typically more flexible/less stable, such as cohabitation. For now, researchers have found no effect of same-sex marriage on the number of different-sex marriages or on the number of divorces (Trandafir, 2014, 2015). However, we wonder whether the legal recognition of same-sex marriage could somehow impact the preferences observed on the different markets. What changes should we expect in the behavior of heterosexuals? And could it be that same-sex couples become more homogamous as same-sex marriage is institutionalized?

Finally, in Becker's theory, a rationale for marriage is the home production complementarities between men and women skills. However, the traditional gains from marriage have diminished for two main reasons. First, the progress in home technology has decreased the

value of domestic production; second, as women took control over their fertility and have become more and more educated, their opportunity cost of staying at home has increased (Stevenson and Wolfers, 2007; Greenwood, Guner, Kocharkov, and Santos, 2016). Despite the decrease in the gains to traditional marriage, the institution of marriage has not disappeared. On the contrary, there has been a high demand for same-sex legal marriage in many developed countries. Stevenson and Wolfers (2007) argue that individuals now look for a mate with whom they “share passions” and the new rationale for marriage is now “consumption complementarities” instead of “production complementarities”. It is also possible that the act of marriage itself is still considered as intrinsically valuable for cultural and social reasons. In any case, this evolution may lead to even higher correlation of traits. Time will tell how these changes will impact macroeconomic outcomes, life quality and social distance among individuals.

1.A. Pooled matching market with sexual orientation

In this appendix, we consider a model with endogenous selection of the partner’s gender. In this model, gender and sexual orientation are observable characteristics among others. This model therefore assumes that all the individuals are pooled into one market, and that the partner’s gender is endogenously chosen and determined based on market characteristics, and in particular, it is subject to trade-offs with other variables. While we include this model for completeness, we regard it as a theoretical construction in the absence of matched data including a measure of sexual of orientation, and we do not present it in the main paper. The model that we use in the main body of the paper operates under the limiting assumption that sexual orientation fully determines the partner’s gender (assumption 3), which is why we assume in the paper that there are three completely segmented markets: lesbian, gay, and heterosexual.

In reality, segmentation may not be perfect, and some people might face trade-offs between matching with someone of their preferred gender rather than an attractive person of the less preferred gender. In addition, some people might be equally attracted by both genders (i.e., bisexual). Assumption 3 implies that we disregard these trade-offs.

Consider a model where individuals are represented by: (i) their economic characteristics x_e ; (ii) their gender x_g , which is a dummy variable equal to 0 if male, 1 if female (this could be extended to a continuous gender spectrum in which case x_g may vary continuously between 0 and 1); and (iii) a measure of sexual orientation x_o which is set so that $x_o = 1$ if the individual is maximally interested in women, and $x_o = 0$ if maximally interested

in men.²⁹ In this case, letting $x = (x_e, x_g, x_o)$, the affinity model will result in an affinity matrix written blockwise as

$$A = \begin{pmatrix} A^{ee} & A^{eg} & A^{eo} \\ * & A^{gg} & A^{go} \\ * & * & A^{oo} \end{pmatrix}$$

where the stars denote terms that are omitted due to the symmetry of A . Several components of A are especially interesting:

- A^{ee} is the classical affinity matrix between the socio-economic and demographic characteristics of the partners, and measures the pairwise assortativeness on education, age, income, race etc. Positive entries of A^{ee} denote complementarity between these characteristics.
- A^{go} denotes the affinity between sexual orientation and gender, which is expected to be positive, by the very definition of x_o .
- A^{gg} denotes the utility penalization of same-sex couples with respect to different-sex ones, which we expect to be negative, in particular due to the relatively higher cost of bearing children for these households, and perhaps also due to social pressure in traditional societies.

In this model, there is a trade-off in surplus between a term that reflects homogamy on socio-economic characteristics (whose strength is determined by A^{ee}), a term that reflects sexual orientation (whose strength is determined by A^{go}), and a term that reflects in particular the higher cost of bearing children (whose strength is determined by A^{gg}).

Note that, in the limit where the affinity term A^{go} between gender and sexual preference is very strong ($A^{go} \rightarrow +\infty$), we get a full segmentation of the markets, where the partner's choice is fully determined by sexual orientation. In this case, we get a sequential choice model where the agents choose in a first stage which market (same-sex or different-sex) to enter, and then choose the remaining partner's characteristics. Because we do not observe sexual orientation, we decided to adopt this limiting case as our framework, thus making the assumption that A^{go} is very large with respect to the other terms. Hence, the partner's gender is fully determined by sexual orientation and own gender; agents select in a first stage into either homo- or heterosexual market, which are fully segmented.

1.B. Descriptive statistics

²⁹Sexual orientation can be measured for instance by the means of the Kinsey scale (see [Sell, 1997](#), for a review), which is a number between 0 (exclusively heterosexual) and 6 (exclusively homosexual). If k is the value of the Kinsey scale, x_o will be set to $k/6$ if the individual is a woman, and to $1 - k/6$ if the individual is a man, which is summarized in the formula $x_o := x_g(k/6) + (1 - x_g)(1 - k/6)$.

Table 1.1. Sample means by sexual orientation

Type of couples	Age	Education	Wage	Hours	Sample size	Share
<i>Married Heterosexuals</i>						87.39%
Men	40.22	12.33	31.34	43.60	124,772	
Women	38.37	12.47	22.84	36.16	124,772	
<i>Unmarried Heterosexuals</i>						11.33%
Men	36.31	11.17	19.79	41.33	16,174	
Women	34.84	11.55	18.45	38.30	16,174	
<i>Homosexuals</i>						
Men	40.00	13.93	35.19	42.71	2,034	0.71%
Women	39.35	13.78	28.44	40.74	1,620	0.57%

Notes: samples are composed of married individuals aged between 25 and 50.

Table 1.2. Distribution of race by sexual orientation

Ethnic	Heterosexual	Gay	Lesbian	All
White	42.7	67.3	63.5	43.0
Black	2.9	2.5	5.2	2.9
Others	0.6	0.8	1.2	0.6
Asian	16.7	8.8	5.8	16.6
Hispanic	37.1	20.7	24.3	36.9
Total	100.0	100.0	100.0	100.0

Notes: samples are composed of married individuals aged between 25 and 50.

Table 1.3. Couples' Pearson correlation coefficients by sexual orientation

Type of couples	Age	Education	Wage	Hours
Heterosexual married couples	0.76	0.71	0.16	-0.12
Heterosexual unmarried couples	0.68	0.64	0.30	0.11
Gay couples	0.56	0.56	0.15	0.03
Lesbian couples	0.66	0.65	0.20	0.08

Table 1.4. Homogamy rates by sexual orientation

	White	Black	Others	Asian	Hispanic		White	Black	Others	Asian	Hispanic
White	1.12	0.79	0.46	0.94	0.67	White	1.28	0.39	0.74	0.70	0.47
Black		12.31	0.00	0.44	0.57	Black		10.67	1.00	0.82	0.53
Others			40.00	2.14	1.21	Others			20.00	0.91	0.87
Asian				3.08	0.33	Asian				8.00	0.13
Hispanic					2.39	Hispanic					2.69

(a) Gays

(b) Lesbians

Men	Women				
	White	Black	Others	Asian	Hispanic
White	1.96	0.32	0.87	0.37	0.28
Black	0.49	24.02	1.38	0.34	0.36
Others	0.84	0.62	60.91	0.40	0.46
Asian	0.15	0.08	0.27	5.08	0.07
Hispanic	0.26	0.16	0.44	0.09	2.32

(c) Heterosexuals

Notes: the homogamy rate is the ratio between the observed number of couples of a certain type and the counterfactual number which should be observed if individuals formed couples randomly.

1.C. Main estimation results

We report our complete estimation results. In this appendix, we present our estimates for the main sample. Table 1.5 presents our estimates of the affinity matrix of each market (table 1.5a for the male same-sex marriage market, table 1.5b for the female same-sex marriage market, table 1.5c for the married different-sex marriage market and table 1.5d

for the unmarried different-sex market). Table 1.6 presents the results of our saliency analysis, i.e., the decomposition of the affinity matrices in orthogonal dimensions. In appendix 1.D, table 1.7 presents our estimates of the affinity matrix when we perform a bipartite estimation of the same-sex marriage market without requiring the affinity matrix to be symmetric. In this case, we define two separate subpopulations to run a bipartite estimation. On one side of the market, we group all those gay individuals that are registered as “householders”, whereas on the other we group their “cohabiting partners”. Table 1.7a displays our estimates for the male same-sex marriage market whereas 1.7b table presents our estimates for the female same-sex marriage market. Finally, in appendix 1.E, table 1.8 presents our estimation results on additional selected samples: 1) childless couples, 2) bi-earner couples, 3) couples living in the metropolitan area of Los Angeles or San Francisco, 4) young couples (25-35 year old). For different-sex married couples, table 1.8c also shows our results for couples with one child only, for couples with three children and more, and for recently married couples with no children. In this table, we do not show all the coefficients of the affinity matrix but only the diagonal coefficients. Each sub-table presents the results for a particular market and each row displays the estimates for a particular selected sub-sample of this market.

Table 1.5. Estimated affinity matrix by marriage market

	Age	Educ.	Wage	Hours	Race		Age	Educ.	Wage	Hours	Race
Age	0.62	-0.06	-0.02	-0.13		Age	0.79	0.04	0.05	-0.10	
	(0.04)	(0.06)	(0.03)	(0.04)			(0.05)	(0.07)	(0.06)	(0.05)	
Education		0.84	0.13	-0.07		Education		1.19	0.20	-0.01	
		(0.09)	(0.06)	(0.06)				(0.12)	(0.10)	(0.07)	
Wage			0.05	-0.07		Wage			0.06	-0.19	
			(0.02)	(0.03)					(0.04)	(0.05)	
Hours				0.12		Hours				0.20	
				(0.04)						(0.05)	
Race					0.62	Race					1.26
					(0.06)						(0.07)
σ					1.26	σ					1.23

(a) Gays (1,017 couples)

(b) Lesbians (810 couples)

	Age	Educ.	Wage	Hours	Race		Age	Educ.	Wage	Hours	Race
Men						Men					
Age	2.17	-0.20	-0.01	-0.03		Age	1.14	-0.06	0.00	-0.06	
	(0.05)	(0.03)	(0.02)	(0.03)			(0.03)	(0.02)	(0.02)	(0.02)	
Education	-0.04	0.82	0.19	-0.04		Education	-0.07	0.66	0.37	0.05	
	(0.03)	(0.03)	(0.03)	(0.03)			(0.02)	(0.02)	(0.04)	(0.02)	
Wage	0.09	0.27	0.01	-0.13		Wage	0.01	0.21	0.05	0.06	
	(0.03)	(0.04)	(0.00)	(0.02)			(0.04)	(0.05)	(0.01)	(0.03)	
Hours	0.06	0.09	-0.09	-0.04		Hours	-0.05	-0.02	0.05	0.09	
	(0.02)	(0.02)	(0.01)	(0.02)			(0.02)	(0.02)	(0.02)	(0.02)	
Race					2.49	Race					1.98
					(0.04)						(0.04)
σ					1.00	σ					1.04

(c) Married heterosexuals (6,228 couples)

(d) Unmarried heterosexuals (5,645 couples)

Notes: the tables display estimates of the affinity matrix A obtained with a sample of couples where both partners are aged between 25 and 50. If the entry A_{ij} is positive and significant, then trait i and j are found to be complements in the marital surplus function. On the contrary, if A_{ij} is negative and significant, i and j are substitutes. Standard errors are in parentheses. Boldfaced estimates are significant at the 5 percent level.

Table 1.6. Estimated indices of attractiveness by marriage market

	I1	I2	Ethnicity
Age	0.11	0.97	
Education	0.94	-0.15	
Wage	0.27	-0.04	
Hours	-0.19	-0.19	
Share of systematic surplus	35%	30%	33%

(a) Gays (1,017 couples)

	I1	I2	Ethnicity
Age	0.98	-0.17	
Education	0.15	0.97	
Wage	0.09	0.19	
Hours	-0.13	-0.03	
Share of systematic surplus	29%	27%	42%

(b) Lesbians (810 couples)

	I1		I2		Ethnicity
	Men	Women	Men	Women	
Age	1.05	1.02	0.04	0.11	
Education	-0.06	-0.11	0.86	0.97	
Wage	0.03	-0.01	0.31	0.18	
Hours	0.02	-0.02	0.09	-0.14	
Share of systematic surplus	36%		16%		48%

(c) Married heterosexuals (6,228 couples)

	I1		I2		Ethnicity
	Men	Women	Men	Women	
Age	0.96	0.95	0.12	0.12	
Education	-0.13	-0.12	0.97	0.93	
Wage	-0.02	-0.04	0.30	0.51	
Hours	-0.04	-0.05	-0.01	0.08	
Share of systematic surplus	33%		20%		47%

(d) Unmarried heterosexuals (5,645 couples)

Notes: Each column displays the estimates of factor loadings explaining the composition of the p -th index of attractiveness \tilde{x}^p and the corresponding share of average systematic surplus $\mathbb{E}_{\hat{\pi}}[\Phi(X, Y)]$ explained by such index (see section 1.5.5). For each market, we present the two indices that explain the largest shares of surplus, as well as the share of surplus explained by ethnicity. Estimates are obtained with a sample of couples where both partners are aged between 25 and 50.

1.D. Bipartite estimation for same-sex couples: head/spouse

Table 1.7. Estimated affinity matrix for same-sex couples with asymmetric roles

Head	Partner				
	Age	Educ.	Wage	Hours	Race
Age	0.60 (0.04)	-0.06 (0.05)	-0.00 (0.03)	-0.17 (0.04)	
Education	0.00 (0.06)	0.80 (0.08)	0.12 (0.05)	-0.02 (0.06)	
Wage	-0.01 (0.04)	0.16 (0.07)	0.06 (0.02)	-0.07 (0.04)	
Hours	-0.05 (0.04)	-0.09 (0.05)	-0.04 (0.03)	0.11 (0.04)	(0.03)
Race					0.65 (0.06)
σ	1.36				

(a) Gays (1,017 couples)

Head	Partner				
	Age	Educ.	Wage	Hours	Race
Age	0.74 (0.05)	0.02 (0.07)	0.06 (0.05)	-0.05 (0.04)	
Education	0.05 (0.07)	1.18 (0.12)	0.18 (0.09)	0.05 (0.07)	(0.04)
Wage	0.02 (0.05)	0.17 (0.10)	0.08 (0.03)	-0.24 (0.05)	
Hours	-0.11 (0.04)	-0.07 (0.07)	-0.15 (0.05)	0.21 (0.04)	
Race					1.18 (0.07)
σ	1.33				

(b) Lesbians (810 couples)

Notes: the tables display estimates of A in a bipartite market where one side is represented by the population of “heads of household” and the other side by the “head’s partners”. Contrarily to the matrices A estimated in 1.5a and 1.5b, now A does not need to be symmetric: symmetry tests can be found in the online appendix. We use a sample of same-sex couples where both partners are aged between 25 and 50. Standard errors are in parentheses. Boldfaced estimates are significant at the 5 percent level.

1.E. Robustness checks

Table 1.8. Summary tables: main findings and robustness checks

	Age	Educ.	Wage	Hours	Race
All	0.62 (0.04)	0.84 (0.09)	0.05 (0.02)	0.12 (0.04)	0.62 (0.06)
Childless	0.56 (0.04)	0.65 (0.09)	0.06 (0.02)	0.18 (0.04)	0.46 (0.07)
Both working	0.64 (0.05)	1.20 (0.12)	0.06 (0.02)	0.29 (0.09)	0.56 (0.07)
Urban	0.60 (0.05)	0.94 (0.13)	0.04 (0.02)	0.16 (0.05)	0.49 (0.08)
25-35 year old	2.48 (0.49)	1.07 (0.28)	0.35 (0.13)	0.01 (0.13)	0.88 (0.16)

(a) Gay couples

	Age	Educ.	Wage	Hours	Race
All	0.79 (0.05)	1.19 (0.12)	0.06 (0.04)	0.20 (0.05)	1.26 (0.07)
Childless	0.76 (0.06)	0.91 (0.15)	0.11 (0.06)	0.34 (0.06)	1.13 (0.09)
Both working	0.85 (0.06)	1.57 (0.16)	0.07 (0.04)	0.20 (0.09)	1.15 (0.08)
Urban	0.87 (0.08)	1.62 (0.21)	-0.00 (0.04)	0.13 (0.07)	1.08 (0.10)
25-35 year old	1.34 (0.34)	1.38 (0.27)	-0.18 (0.31)	0.43 (0.10)	1.33 (0.15)

(b) Lesbian couples

	Age	Educ.	Wage	Hours	Race
All	2.17 (0.05)	0.82 (0.03)	0.01 (0.00)	-0.04 (0.02)	2.49 (0.04)
Childless	1.88 (0.04)	1.01 (0.04)	0.07 (0.01)	0.12 (0.02)	2.10 (0.04)
One child	2.14 (0.05)	0.81 (0.03)	0.04 (0.01)	0.02 (0.02)	2.38 (0.04)
Three children	2.38 (0.05)	0.75 (0.02)	0.02 (0.00)	-0.02 (0.02)	2.69 (0.04)
Newlyweds, childless	1.40 (0.05)	1.37 (0.09)	0.02 (0.01)	0.16 (0.03)	1.81 (0.05)
Both working	2.44 (0.05)	0.94 (0.04)	0.04 (0.01)	0.20 (0.04)	2.36 (0.04)
Urban	2.18 (0.05)	0.86 (0.03)	0.02 (0.00)	-0.04 (0.02)	2.59 (0.04)
25-35 year old	6.33 (0.17)	1.42 (0.05)	0.07 (0.02)	0.06 (0.03)	3.09 (0.05)

(c) Married couples

	Age	Educ.	Wage	Hours	Race
All	1.14 (0.03)	0.66 (0.02)	0.05 (0.01)	0.09 (0.02)	1.98 (0.04)
Childless	1.11 (0.03)	0.90 (0.04)	0.04 (0.01)	0.24 (0.02)	1.47 (0.04)
Both working	1.19 (0.03)	0.70 (0.03)	0.18 (0.02)	0.47 (0.04)	1.83 (0.03)
Urban	1.15 (0.03)	0.61 (0.02)	0.02 (0.0)	0.08 (0.02)	2.00 (0.04)
25-35 year old	3.58 (0.11)	1.08 (0.04)	0.15 (0.03)	0.19 (0.03)	2.45 (0.04)

(d) Unmarried couples

Notes: Each row displays the estimates of the diagonal coefficients of the affinity matrix A obtained with a given sample. The first row (“All”) refers to our benchmark results already presented in table 1.5. The other rows refer to alternative subsamples used to conduct our auxiliary estimations. Complete tables with all entries of A are available in the online appendix. Standard errors are in parentheses. Boldfaced estimates are significant at the 5 percent level.

Bibliography

- ALDÉN, L., L. EDLUND, M. HAMMARSTEDT, AND M. MUELLER-SMITH (2015): “Effect of registered partnership on labor earnings and fertility for same-sex couples: evidence from Swedish register data,” *Demography*, 52(4), 1243–1268.
- ALLEN, D. W., AND S. E. LU (2017): “Matching, marriage, and children: differences across sexual orientations,” *Review of Economics of the Household*, 15(2), 527–547.
- ANDERSSON, G., T. NOACK, A. SEIERSTAD, AND H. WEEDON-FEKJÆR (2006): “The demographics of same-sex marriages in Norway and Sweden,” *Demography*, 43(1), 79–98.
- ANTECOL, H., AND M. D. STEINBERGER (2013): “Labor Supply Differences between Married Heterosexual Women and Partnered Lesbians: A Semi-Parametric Decomposition Approach,” *Economic Inquiry*, 51(1), 783–805.
- BECKER, G. (1973): “A theory of marriage: Part I,” *Journal of Political Economy*, 81(4).
——— (1991): *A Treatise on the Family*. Harvard University Press.
- BLACK, D., G. GATES, S. SANDERS, AND L. TAYLOR (2002): “Why do gay men live in San Francisco?,” *Journal of Urban Economics*, 51(1), 54–76.
- BLACK, D. A., S. G. SANDERS, AND L. J. TAYLOR (2007): “The economics of lesbian and gay families,” *The Journal of Economic Perspectives*, 21(2), 53–70.
- BLUMSTEIN, P., P. SCHWARTZ, ET AL. (1983): *American couples: Money, work, sex*. Morrow New York.
- BRIEN, M. J., L. A. LILLARD, AND S. STERN (2006): “Cohabitation, marriage, and divorce in a model of match quality,” *International Economic Review*, 47(2), 451–494.
- CANDELON, B., AND A. DUPUY (2015): “Hierarchical organization and performance inequality: evidence from professional cycling,” *International Economic Review*, 56(4), 1207–1236.
- CHIAPPORI, P. A., A. GALICHON, AND B. SALANIÉ (forthcoming): “On Human Capital and Team Stability,” *Journal of Human Capital*.
- CHIAPPORI, P.-A., S. OREFFICE, AND C. QUINTANA-DOMEQUE (2012): “Fatter attraction: anthropometric and socioeconomic matching on the marriage market,” *Journal of Political Economy*, 120(4), 659–695.
- CHIAPPORI, P.-A., B. SALANIÉ, AND Y. WEISS (2017): “Partner Choice, Investment in Children, and the Marital College Premium,” *American Economic Review*, 107(8), 2109–67.
- CHOO, E., AND A. SIOW (2006): “Who Marries Whom and Why,” *Journal of Political Economy*, 114, 175–201.
- DAGSVIK, J. K. (1994): “Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes,” *Econometrica: Journal of the Econometric Society*, pp. 1179–1205.
- (2000): “Aggregation in matching markets,” *International Economic Review*, 41(1), 27–58.

- DIAMOND, L. M. (2008a): “Female bisexuality from adolescence to adulthood: results from a 10-year longitudinal study,” *Developmental psychology*, 44(1), 5.
- (2008b): *Sexual fluidity*. Harvard University Press.
- DILLON, F. R., R. L. WORTHINGTON, AND B. MORADI (2011): “Sexual identity as a universal process,” in *Handbook of identity theory and research*, ed. by S. J. Schwartz, K. Luyckx, and V. L. Vignoles, pp. 649–670. Springer.
- DUPUY, A., AND A. GALICHON (2014): “Personality Traits and the Marriage Market,” *Journal of Political Economy*, 122(6), 1271–1319.
- EIKA, L., M. MOGSTAD, AND B. ZAFAR (forthcoming): “Educational Assortative Mating and Household Income Inequality,” Discussion paper.
- ESKRIDGE JR, W. N. (1993): “A history of same-sex marriage,” *Virginia Law Review*, pp. 1419–1513.
- FOX, J. T. (2010): “Identification in matching games,” *Quantitative Economics*, 1(2), 203–254.
- (2018): “Estimating matching games with transfers,” *Quantitative Economics*, 9(1), 1–38.
- GALICHON, A., AND B. SALANIÉ (2014): “Cupid’s invisible hand: Social surplus and identification in matching models,” Discussion paper, SSRN.
- GATES, G. J. (2010): “Same-sex couples in US Census Bureau Data: Who gets counted and why,” Discussion paper, the Williams Institute.
- GEMICI, A., AND S. LAUFER (2011): “Marriage and cohabitation,” *New York University, mimeo*.
- GOLDIN, C., L. F. KATZ, AND I. KUZIEMKO (2006): “The Homecoming of American College Women: The Reversal of the College Gender Gap,” *The Journal of Economic Perspectives*, pp. 133–156.
- GOUSSÉ, M., N. JACQUEMET, AND J.-M. ROBIN (2017): “Marriage, labor supply, and home production,” *Econometrica*, 85(6), 1873–1919.
- GREENWOOD, J., N. GUNER, G. KOCHARKOV, AND C. SANTOS (2016): “Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation,” *American Economic Journal: Macroeconomics*, 8(1), 1–41.
- HARRY, J. (1984): *Gay couples*. Praeger Publishers.
- HAYES, A. F. (1995): “Age preferences for same-and opposite-sex partners,” *The Journal of social psychology*, 135(2), 125–133.
- HECKMAN, J. (2007): “Notes on Koopmans and Beckmann’s ” Assignment Problems and the Location”,” Lecture notes, University of Chicago.
- HINES, M. (2011): “Gender development and the human brain,” *Annual review of neuroscience*, 34, 69–88.
- HITSCH, G. J., A. HORTAÇSU, AND D. ARIELY (2010): “Matching and sorting in online dating,” *American Economic Review*, 100(1), 130–63.

- JANNINI, E. A., R. BLANCHARD, A. CAMPERIO-CIANI, AND J. BANCROFT (2010): “Male homosexuality: Nature or culture?,” *The Journal of Sexual Medicine*, 7(10), 3245–3253.
- JEPSEN, C., AND L. K. JEPSEN (2015): “Labor-Market Specialization within Same-Sex and Difference-Sex Couples,” *Industrial Relations: A Journal of Economy and Society*, 54(1), 109–130.
- JEPSEN, L. K., AND C. A. JEPSEN (2002): “An Empirical Analysis of the Matching Patterns of Same-Sex and Opposite-Sex Couples,” *Demography*, 39(3), 435–453.
- KALMIJN, M. (1998): “Intermarriage and homogamy: Causes, patterns, trends,” *Annual Review of Sociology*, 24(1), 395–421.
- KENRICK, D. T., AND R. C. KEEFE (1992): “Age preferences in mates reflect sex differences in human reproductive strategies,” *Behavioral and Brain Sciences*, 15(01), 75–91.
- KURDEK, L. A., AND J. P. SCHMITT (1987): “Partner homogamy in married, heterosexual cohabiting, gay, and lesbian couples,” *Journal of Sex Research*, 23(2), 212–232.
- LOW, C. (2013): “Pricing the Biological Clock: Reproductive Capital on the US Marriage Market,” Discussion paper, Mimeo.
- MEIER, A., K. E. HULL, AND T. A. ORTYL (2009): “Young adult relationship values at the intersection of gender and sexuality,” *Journal of Marriage and Family*, 71(3), 510–525.
- MENZEL, K. (2015): “Large Matching Markets as Two-sided Demand Systems,” *Econometrica*, 83, 306–337.
- MOURIFIÉ, I. Y., AND A. SIOW (2014): “Cohabitation vs Marriage: Marriage matching with peer effects,” Discussion paper, University of Toronto.
- OREFFICE, S. (2011): “Sexual orientation and household decision making.: Same-sex couples’ balance of power and labor supply choices,” *Labour Economics*, 18(2), 145–158.
- PESKI, M. (2017): “Large roommate problem with non-transferable random utility,” *Journal of Economic Theory*, 168, 432–471.
- PETRONGOLO, B., AND C. A. PISSARIDES (2001): “Looking into the black box: A survey of the matching function,” *Journal of Economic literature*, 39(2), 390–431.
- ROSENFELD, M. J., AND B.-S. KIM (2005): “The independence of young adults and the rise of interracial and same-sex unions,” *American Sociological Review*, 70(4), 541–562.
- SCHWARTZ, C. R., AND N. L. GRAF (2009): “Assortative Matching Among Same-Sex and Different-Sex Couples in the United States, 1990–2000,” *Demographic research*, 21, 843.
- SELL, R. L. (1997): “Defining and measuring sexual orientation: A review,” *Archives of sexual behavior*, 26(6), 643–658.
- SHIMER, R., AND L. SMITH (2000): “Assortative matching and search,” *Econometrica*, 68(2), 343–369.

- STEVENSON, B. (2008): “The Evolution of the American Family: An Economic Interpretation,” *American Journal of Family Law*, 22(3), 141–147.
- STEVENSON, B., AND J. WOLFERS (2007): “Marriage and Divorce: Changes and Their Driving Forces,” *The Journal of Economic Perspectives*, pp. 27–52.
- SULLIVAN, A. (2009): *Same-sex marriage: Pro and con*. Random House LLC.
- TRANDAFIR, M. (2014): “The effect of same-sex marriage laws on different-sex marriage: Evidence from the Netherlands,” *Demography*, 51(1), 317–340.
- (2015): “Legal recognition of same-sex couples and family formation.,” *Demography*, 52(1), 113–151.
- VERBAKEL, E., AND M. KALMIJN (2014): “Assortative Mating Among Dutch Married and Cohabiting Same-Sex and Different-Sex Couples,” *Journal of Marriage and Family*, 76(1), 1–12.
- VIRGILE, M. (2011): “Measurement error in the relationship status of same-sex couples in the 2009 American Community Survey,” Discussion paper, ACS11-RER-10, US Government Printing Office, Washington, DC.
- WEISS, Y., J. YI, AND J. ZHANG (2013): “Hypergammy, Cross-Boundary Marriages, and Family Behavior,” Discussion paper, IZA Discussion Paper.

CHAPTER 2

The role of evolving marital preferences in growing income inequality¹

Abstract

In this paper, we describe mating patterns in the United States from 1964 to 2017 and measure the impact of changes in marital preferences on between-household income inequality. We rely on the recent literature on the econometrics of matching models to estimate complementarity parameters of the household production function. Our structural approach allows us to measure sorting on multiple dimensions and to effectively disentangle changes in marital preferences and in demographics, addressing concerns that affect results from existing literature. We answer the following questions: *has assortativeness increased over time? Along which dimensions? To what extent the shifts in marital preferences can explain inequality trends?* We find that, after controlling for other observables, assortative mating in education has become stronger. Moreover, if mating patterns had not changed since 1971, the 2017 Gini coefficient between married households would be lower by 6%. We conclude that about 25% of the increase in between-household inequality is due to changes in marital preferences. Increased assortativeness in education positively contributes to the inequality rise, but only modestly.

2.1. Introduction

The study of mating patterns, and especially assortativeness, traces back to the seminal work of [Becker \(1973, 1974, 1991\)](#). Becker's earliest model of a competitive marriage market aims at rationalizing both household specialization and the homogamy observed in the data with respect to several non-labor market traits (e.g. education, ethnicity, religion). Becker points at the structure of the household production function to explain marriage patterns: complementarity between inputs leads to optimal positive assortative mating, whereas substitutability leads to negative assortative mating.

In light of such observations on Becker's work, studying marriage patterns is primarily insightful because it reveals much about intra-household dynamics. Differences in mating

¹Chapter coauthored with Simon Weber.

dynamics over time and space may be the result of transformations in the institution of the family, labor market conditions, available household-production technology, gender roles, etc. For instance, one could wonder whether Becker’s observation that we should expect a negative association between spouses’ wage rates due to household specialization still applies to modern families despite the improvements in home technology and closing gender wage gap.² Changes in the cultural and legal framework also matter for the evolution of marital preferences, due to their influence on marriage flows and on the allocation of resources across and within couples.

In recent years, marital sorting has become the object of increasing attention because of its relationship with growing inequalities between households. Researchers have focused on the relationship between marriage patterns, between-household income inequality and long-run economic outcomes (e.g. [Burtless, 1999](#); [Fernández et al., 2005](#)). The compelling research question is whether stronger assortativeness with respect to some crucial dimensions - notably, education - is associated with higher inequality.

The aim of this paper is to build a connection between changes in the structure of marital gains and the increasing income inequality observed in the United States. We address the following questions: has assortative mating increased over time? And, if yes, along which dimensions? What is the impact of shifts in marital preferences on household income inequality? The framework we adopt follows [Choo and Siow \(2006\)](#) and [Galichon and Salanié \(2015\)](#)’s observation that joint marital surplus can be identified with data on matches in a static, competitive matching framework. We employ the recent estimation technique proposed by [Dupuy and Galichon \(2014\)](#) and estimate the degree of complementarity and substitutability between the spouses’ traits. Such estimates stand as our measures of the strength of marital sorting.

This structural approach allows us to contribute to the literature on sorting and inequality by overcoming some limitations affecting studies based on standard measures of assortativeness, such as correlation coefficients, homogamy rates, frequency tables, and so on. Disentangling changes in marital preferences and demographics is crucial because of important changes in the marginal distributions of people’s traits in the United States during the last decades (e.g., aging of the population, overall increase in schooling attainment, closing of the gender wage gap, and reversal of the gender gap in higher education). In addition, our analysis is not limited to educational assortativeness: the multidimensional matching model of [Dupuy and Galichon \(2014\)](#) provides tools to study complementarity on education, as well as interactions between other socio-economic traits. Following this new approach, we rediscuss the findings of several key papers in the marriage literature,

²The survey of [Stevenson and Wolfers \(2007\)](#) keeps track of the changes that the institution of the family has gone through in recent decades, and presents several significant research questions that need to be answered.

such as [Fernández et al. \(2005\)](#), [Schwartz and Mare \(2005\)](#) and the recent [Greenwood et al. \(2014\)](#), [Eika et al. \(forthcoming\)](#) and [Greenwood et al. \(2016\)](#).

The theoretical framework of [Dupuy and Galichon \(2014\)](#) is grounded on Transferable Utility models and logit formalism, and extends the seminal matching model of [Choo and Siow \(2006\)](#) to the multidimensional and continuous case. Agents are fully informed about potential partners' characteristics, but the econometrician only observes traits $x \in X$ and $y \in Y$, respectively for men and women, where X and Y are continuous and multidimensional. The empirical strategy relies on a bilinear parametrization for the systematic marriage surplus function, i.e. $\Phi(x, y) = x' Ay$. It follows that we can measure the degree of complementarity or substitutability by estimating the *marital preference parameters*, i.e. the elements of the *affinity matrix* A , since $\partial^2 \Phi / \partial x_j \partial y_k = A_{jk}$. These will be our measures of assortativeness. In addition, after estimating A , we recover the optimal probability distribution of matches $\pi^A(x, y)$, which, in other words, is the simulated joint frequency table of partners' types at equilibrium. The latter depends both on the structure of preferences given by A and the marginal distributions of observable types $f(x)$ and $g(y)$: operating on the parameters A , we can compute the predicted distribution of couples' traits under counterfactual preferences. For instance, we can artificially increase the value of one parameter of A , say the strength of assortative mating on education, and check how the distribution of partners' types $\pi^A(x, y)$ changes at the new (counterfactual) marriage market equilibrium.

In practice, we estimate marital preference parameters for the United States over the period 1964-2017 with Current Population Survey data to keep track of sorting dynamics through the analysis of preferences. We consider the following observable variables: age, education, hourly wage, hours worked, and ethnic background. We subsequently use the marriage patterns predicted by the model - the optimal matching function $\pi^A(x, y)$ - to construct counterfactual household income distributions. To do so, we substitute the actual preferences measured for a given wave with counterfactual preferences measured for a different wave. This means that we provide a prediction of how people would have sorted into married couples in a given year if their marital preferences had been equal to those of another cohort (e.g. to their parents' or grandparents'). In this way, we study the contribution of changes in marital preferences to the observed marriage patterns and to the evolution of inequality in several illustrative examples.

To the best of our knowledge, this is the first attempt to analyze the evolution of marital preferences in the United States by means of structural estimation techniques in a multidimensional matching framework.³ We hereby provide a complete analysis of assortativeness along multiple observable socio-economic traits, track changes in sorting

³Two related works are those by [Chiappori et al. \(2017\)](#) and [Greenwood et al. \(2016\)](#), but both focus on educational sorting. Their findings are discussed in Section 2.5.

patterns over time and assess to what extent they can explain the rise in between-household income inequality in the last decades.

The paper is organized as follows. Section 2.2 provides a brief literature review, while Section 2.3 introduces the theoretical framework. In Section 2.4, we describe CPS data and our sample selection criteria. Then, we present and discuss our results: in Section 2.5, the trends in marital preferences, while in Section 2.6 the counterfactual analysis of inequality. Section 2.7 concludes.

2.2. Previous findings

2.2.1. Evolution of mating patterns. A crucial question that the literature has tried to answer in different ways is whether assortativeness has increased over time. The demographic and sociological literature often makes use of log-linear models to explain mating patterns and measure assortativeness. Log-linear models for contingency tables help “specify how the size of a cell count depends on the levels of the categorical variables for that cell” (Agresti, 2013, Chapter 9). Several papers relying on this methodology focus on assortativeness on education: the contingency table of size $I \times I$ tells the frequency of couples by partners’ education ij , with $i, j \in \{1, \dots, I\}$ being the individual schooling level. If matching were random, the following regression would exhibit a good fit

$$\log \mu_{ij} = \lambda + \lambda_i^M + \lambda_j^W \quad (2.2.1)$$

where μ_{ij} is the frequency of a couple with education ij , λ_i^M is the vector of men’s educational level effects and λ_j^W is the vector of women’s educational level effects. Under random matching, marginal distributions are sufficient to explain the entries of the contingency tables. Nonetheless, if matching is not random, then one needs to include other regressors to explain the couples’ joint distribution. “Homogamy models” contain an additional regressor measuring the impact of educational homogamy on the log-joint frequency $\log \mu_{ij}$ (e.g. Johnson, 1980; Kalmijn, 1991b; Schwartz and Mare, 2005). “Crossing models”, instead, contain additional regressors measuring the impact of crossing an educational barrier (e.g. a college-graduate marrying a dropout, see Mare, 1991; Smits et al., 1998; Schwartz and Mare, 2005). Log-linear models can be rewritten as multinomial choice models (see Agresti, 2013; Schwartz and Graf, 2009), which in turn are surprisingly close in spirit to the matching model class of Choo and Siow (2006). In the equivalent multinomial logit model, the categorical response variable would be the wife’s (or the husband’s) education to represent the choice of the husband (or the wife’s) conditional on his (her) schooling level. However, a basic choice model of this kind does not take into account that men and women actually seek a partner in a competitive environment: the choice of one agent affects the pool of partners available for other agents. As a consequence, it is not possible to interpret the coefficients as the “true” preference parameters. In the

structural framework proposed by [Choo and Siow \(2006\)](#) and [Galichon and Salanié \(2015\)](#), it is instead possible to estimate the parameters of the model so that the matching market is indeed at equilibrium. In these equilibrium models, every agent’s choice is constrained by the choices of other “competitors” and the market must clear, i.e., the sum of singles and married must be equal to the total number of individuals by type and sex.

Several studies apply log-linear models or closely related ones to study changes in educational assortativeness in marriage patterns in the United States. Most agree that educational assortative mating strengthened in the second part of 20th century ([Mare, 1991](#); [Kalmijn, 1991a,b](#); [Qian and Preston, 1993](#)) and the first decade of the 21st ([Schwartz and Mare, 2005](#)), although some other studies argue that educational homogamy stayed constant or declined: for instance, [Fu and Heaton \(2008\)](#) observe a decline between 1980 and 2000, while [Liu and Lu \(2006\)](#) maintain that the intensity of educational homogamy increased from 1960 to 1980 but then started decreasing. Interestingly, most papers also agree that one of the strongest trend is the increase in the frequency of marriages between highly educated individuals. Several papers use log-linear models to explore other matching dimensions, sometimes in multidimensional frameworks, although the number of variables stays low (2 or 3 typically) because of methodological limitations. [Johnson \(1980\)](#) and [Kalmijn \(1991a\)](#) analyze religion, [Schoen and Wooldredge \(1989\)](#) and [Fu and Heaton \(2008\)](#) ethnicity, [Qian and Preston \(1993\)](#) age, [Kalmijn \(1991b\)](#) and [Blackwell \(1998\)](#) parents’ education, [Stevens and Schoen \(1988\)](#) language spoken. Some empirical findings on assortativeness in the United States are particularly interesting since they can be compared with ours. [Qian and Preston \(1993\)](#) find that homogamy with respect to age increased (from 1972 to 1987); [Fu and Heaton \(2008\)](#) find that racial homogamy decreased (from 1980 to 2000).

In the economic literature, some analyses of mating patterns rely on simple descriptive statistics: for instance, [Fryer \(2007\)](#) uses the probabilities of crossing racial barriers to describe the patterns of racial intermarriage in the United States and explore the possible driving forces behind the trends. Other researchers assess the strength of educational assortativeness through the comparison with counterfactual distributions. The simplest indicators of this kind are “homogamy rates” which are the ratios between the actual frequency of a couples’ joint education and the counterfactual frequency computed under random matching. Contingency tables to compare actual and counterfactual joint distributions are similar (if not identical) to homogamy rates (e.g. [Greenwood et al., 2014](#)). Another possibility is to compare the actual distribution to the counterfactual under perfect positive assortative mating (e.g. [Liu and Lu, 2006](#)). While generally insightful, homogamy rates and similar measures are not suitable for comparisons across different populations and even across different categories within the same population. The *size* of the homogamy rate is hardly comparable when the marginals become smaller. Hence, it is hard to set the comparison between homogamy for PhD graduates, who represent

a small share of the population, and high school diplomas, who represent a wide share. In consequence, researchers opt for aggregate measures of assortativeness that take into account the different size that each category has in the population (e.g. [Greenwood et al., 2014](#); [Eika et al., forthcoming](#)). Using such measures based on homogamy rates, [Eika et al. \(forthcoming\)](#) conclude that marital sorting in the United States on education has slightly increased over the period 1980-2007. The findings of [Greenwood et al. \(2014\)](#) are similar: relying on several measures, some of which based on homogamy rates, they find that assortativeness on education has increased in the period 1960-2005.

2.2.2. Assortativeness and inequality. Another crucial question is whether changes in mating patterns can partly explain the trend of income inequality between households. Many authors are concerned with the possibility that more assortativeness on socio-economic characteristics - particularly on education - can lead to higher household income inequality. Since education is a primary dimension of assortativeness, and since highly educated individuals typically have higher income, more educational homogamy implies that high-income individuals will marry with each other more and more frequently. Nevertheless, it is not straightforward to disentangle the effect of changes in marital preferences from the shifts in the marginal distributions. This is particularly relevant because of the closing of the educational gap between men and women in the last decades and women's increased participation to the labor force.

The landmark contributions by [Fernández and Rogerson \(2001\)](#) and [Fernández et al. \(2005\)](#) make an attempt to model the trends of household inequality in order to shed some light on the role played by sorting, fertility and children's education. [Fernández and Rogerson \(2001\)](#) set a model in which individuals are either skilled or unskilled and marry more or less frequently with partners of the same educational level according to an exogenous parameter accounting for the degree of homogamy on the marriage market. Since the children of highly educated families will be more likely to go to college, mating patterns are crucial in order to explain the steady state level of inequality. [Fernández et al. \(2005\)](#) introduce a simple two-round matching model in order to endogenize the strength of sorting on education. They find that, at steady state, a higher degree of sorting - measured as the correlation between partners' income - is associated with higher income inequality.⁴ Both papers argue that educational assortativeness exacerbates inequality in the long run, in disagreement with [Kremer \(1997\)](#), who states that sorting has a negligible impact on steady state inequality. Although the structural approach of these models is extremely insightful to understand through which channels mating patterns may influence inequality in the long run, we believe that their conclusions might - to some extent - depend on their specific measure of educational assortativeness. In particular, [Fernández et al. \(2005\)](#) show that the Pearson correlation coefficient between partners' education

⁴Both [Fernández and Rogerson \(2001\)](#) and [Fernández et al. \(2005\)](#) use the skill premium as a measure of inequality.

correctly measures the degree of assortativeness. However, this conclusion can be reached only under the restrictive assumptions necessary for their two-round matching model. Indeed, in most alternative matching models, a change in the correlation rate may well be due to a change in *marital preferences* as well as to a *shift in the marginals*. Hence, since a higher correlation rate does not necessarily imply more assortativeness, we propose to relate alternative measures of assortativeness to income inequality in order to check whether their conclusions are robust.

As previously mentioned, [Greenwood et al. \(2016\)](#) set up a model of educational choice, marriage and the household, and estimate its steady-state. With respect to the papers mentioned above, the focus is now more on household technology and changes in the wage distribution rather than intergenerational transmission. The authors run a number of counterfactual experiments that help understand what forces contributed to the rise of inequality. In particular, they assess that changes in the wage structure alone explain 39% of such increase. They subsequently stress that changes in marriage patterns account for 18.6% of the increase, which grows to 35.6% when allowing households to adjust their labor supply. In the present paper, we also disentangle changes in the wage distribution from transformations to the structure of marital gains, while we also control for changes in the marginal distribution of other observables (e.g. race and ethnicity). On the other hand, [Greenwood et al. \(2016\)](#) make explicit assumptions on household behavior and their model insightfully predicts how households adjust their labor supply. In this way, they separately assess the effects of changes in home technology and in taste for educational homogamy on income inequality. We compare our empirical findings to theirs in Section 2.6.

Beside the above-mentioned papers, most research focuses on the empirics in the hope of assessing the impact of changes in marital preferences on income inequality in the United States correctly. Measuring the strength of educational assortativeness is not straightforward and several approaches have been tried out. The work by [Burtless \(1999\)](#) is an early example of counterfactual analysis of inequality. In order to assess the degree of inequality that we would observe in 1996 if matching patterns did not change since 1979, Burtless shuffles the observed married couples in 1996 and reassigns spouses as follows: if the man whose income had rank r married a woman with rank s in 1979, the man with rank r in 1996 is assigned to the woman with rank s from the same year. [Cancian and Reed \(1998\)](#) and [Western et al. \(2008\)](#) suggest using decomposition methods on the changes in the variance of household income. The methodology consists in dividing the household population into groups according to certain characteristics (e.g. age, education, children) and then studying the trends of income variance within and between groups.

[Schwartz \(2010\)](#) focuses on marital preferences and is thus more closely related to our analysis. She uses the log-linear models explained in Section 2.2.1 to build counterfactual

distributions of partners' income.⁵ The author concludes that inequality would have been lower without the shifts in income assortativeness.⁶

The works of [Greenwood et al. \(2014\)](#) and [Eika et al. \(forthcoming\)](#) also aim to assess the impact of changes in educational assortativeness on inequality.⁷ Using contingency tables, [Greenwood et al. \(2014\)](#) show that, under random matching, the counterfactual Gini coefficient in 2005 for United States would be lower than the actual (about 2.3% less). In addition, using standardized contingency tables with several controls (e.g. children, participation in the labor force), they assess that, had sorting patterns been constant since 1960, the 2005 Gini coefficient would be almost unaffected (only about 0.3% less). [Eika et al. \(forthcoming\)](#) conduct a similar analysis to study the trends of household income inequality in the United States between 1980 and 2007. They employ a methodology which consists in building counterfactuals by combining the partners' joint distribution of schooling attainments from a given year to the conditional distribution of income given the educational level from another year. They conclude that, had returns to schooling not changed since 1980, 2007 household income inequality would have been much lower (about 23% less). In addition, the authors also remark that, without the overall increase in schooling attainments at individual level, 2007 inequality would be even higher. Finally, they assess that, had 1980 marital preferences been the same as in 2007, we would have not observed any relevant difference in household income inequality: their findings are thus consistent with those of [Greenwood et al. \(2014\)](#), although the time lapse considered is different.

2.3. Theoretical framework

[Dupuy and Galichon \(2014, hereafter DG\)](#) extend the setting of [Choo and Siow \(2006\)](#) and [Galichon and Salanié \(2015\)](#) to the multidimensional and continuous case. Here, we closely follow the methodology of DG. Here we briefly recall the theoretical framework and the estimation technique.

2.3.1. Matching model. In this frictionless Transferable Utility framework, men and women are characterized by a vector of characteristics $x \in \mathcal{X}$ for men, and $y \in \mathcal{Y}$ for women. Note that, with a large set of continuous variables, every individual is virtually unique in his (her) *observable type* given by x (y). A *matching* is a probability distribution

⁵The methodology consists in finding a log-linear model with good fit to explain a contingency table with the distribution of income by percentile (plus one category containing zero-income observations). Then one can compute predicted frequencies after removing certain regressors to reproduce counterfactual situations.

⁶[Schwartz \(2010\)](#) uses the ratios between the median income of the top 20% households (high class) over the median income of the middle 60% (medium class) or the median income of the top 20% (low class) as measures of inequality.

⁷As for [Greenwood et al. \(2014\)](#), we hereby refer to the revised findings published in the corrigendum.

that tells the odds of a couple with observable types x and y to be matched. When a man x and a woman y match, they receive *systematic utility shares* U and V respectively, which both depend on the combination of observable types (x, y) only. In addition, a man of type x experiences a random *sympathy shock* ε^k that is individual-specific to the potential partner k of type y^k . Hence, the two components being additive, the man's payoff from a match with a woman k of type y is given by $U(x, y^k) + \frac{\sigma}{2}\varepsilon^k$, where the scalar σ measures the relevance of the unobservable component. Women's payoff can be written in an analogous way.

When the sympathy shock is of Gumbel type, the setting is completely analogous to [Choo and Siow \(2006\)](#). However, DG suggest assuming that each man chooses his partner within a set of infinite but countable "acquaintances", each with characteristics (y^k, ε^k) over the space $\mathcal{Y} \times \mathbb{R}$: such set is the enumeration of a Poisson process with intensity $dy \times e^{-\varepsilon}d\varepsilon$, which leads us to a continuous logit framework. Under this assumption, the shock ε^k is independent from the observables. Every man solves the following problem

$$\max_k \left\{ U(x, y^k) + \frac{\sigma}{2}\varepsilon^k \right\} \quad (2.3.1)$$

and so do women with due changes in notation.

DG show that it is possible to recover the optimal matching $\pi(x, y)$ and the equilibrium shares $U(x, y)$ and $V(x, y)$ so that the scarcity constraints hold and $\Phi(x, y) \equiv U(x, y) + V(x, y)$, as implied by the Transferable Utility assumption, where $\Phi(x, y)$ denotes the *systematic surplus*. Provided two functions $a(x)$ and $b(y)$ so that $\pi(x, y)$ is *feasible* - the sum of married individuals of a given type does not exceed their initial number - the equilibrium is thus fully characterized by:

- (1) the optimal matching function $\pi(x, y)$, which tells the probability of matching (equivalently, the relative frequency at equilibrium) for a couple with observables (x, y) :

$$\pi(x, y) = \exp \left(\frac{\Phi(x, y) - a(x) - b(y)}{\sigma} \right). \quad (2.3.2)$$

- (2) the shares of systematic surplus at equilibrium for each couple with observables (x, y) :

$$U(x, y) = \frac{\Phi(x, y) + a(x) - b(y)}{2} \quad (2.3.3)$$

$$V(x, y) = \frac{\Phi(x, y) + b(y) - a(x)}{2} \quad (2.3.4)$$

so that $U(x, y) + V(x, y)$ gives the total systematic surplus at equilibrium, i.e., $\Phi(x, y)$.

2.3.2. Specification. In this paper, we consider the following parametrization of the systematic surplus, introduced by [Ciscato et al. \(forthcoming\)](#):

$$\Phi(x, y) = x' Ay + \sum_{i, j \in \{1, \dots, O\}} x_i a_{ij} y_j + \sum_{i \in \{O+1, \dots, O+U\}} \lambda_i \mathbb{1}[x_i = y_i]. \quad (2.3.5)$$

where the first O observable variables are *ordered* and the last U are *unordered*. Examples of ordered variables are age, education and wage, whereas ethnicity and working sector are unordered. Note that transformations of raw variables, such as polynomials, logarithms and ranks, could be added as additional controls.

Our main specification implies that the matrix of parameters A - called *affinity matrix* - looks as follows:

$$A = \begin{bmatrix} \tilde{A} & 0 \\ 0 & \Lambda \end{bmatrix}. \quad (2.3.6)$$

The $O \times O$ entries of the submatrix \tilde{A} determine whether the (ordered) variables are complementary or substitutes, as well as the intensity of the affinity (or repulsion) between the two inputs. The elements of the diagonal submatrix Λ tell us whether homogamy with respect to one of the unordered variables results in an increase rather than in a decrease of the systematic surplus. All the other elements of the matrix are constrained to zero.

2.3.3. Estimation. To compute equilibrium quantities, we solve for $a(x)$ and $b(y)$ enforcing the market scarcity constraints through an Iterative Projection Fitting Procedure for given parameters A and σ . Hence, note that, according to the crucial result of [Shapley and Shubik \(1971\)](#), the equilibrium matching of a decentralised matching market is also the one that maximises social gain. We define the function $\mathcal{W}(A, \sigma)$ as follows:

$$\mathcal{W}(A, \sigma) \equiv \max_{\pi \in \mathcal{M}} \{E_{\pi}[x' Ay] - \sigma E_{\pi}[\log \pi(x, y)]\} \quad (2.3.7)$$

where \mathcal{M} is the set of feasible matchings and where expected values with subscript π are taken with respect to the optimal matching probabilities.

DG set the following convex optimisation problem in order to estimate the matrix $B = A/\sigma$:

$$\min_B \mathcal{W}(B, 1) - E_{\hat{\pi}}[x' By] \quad (2.3.8)$$

where the expected value with subscript $\hat{\pi}$ is taken with respect to the relative frequencies observed in the data. The First Order Conditions of the problem imply that we are matching the co-moments of men's and women's characteristics predicted by the model with the corresponding empirical co-moments observed in the data. In practice, we are computing B so that the following holds

$$E_{\pi}[X_i Y_j] = E_{\hat{\pi}}[X_i Y_j] \quad (2.3.9)$$

for each couple (i, j) of ordered characteristics. Similarly, B must be such that the following holds

$$E_\pi[\mathbf{1}[X_i = Y_i]] = E_{\hat{\pi}}[\mathbf{1}[X_i = Y_i]] \quad (2.3.10)$$

for each unordered characteristic i .

2.3.4. Identification with multiple markets. One drawback of the original model of DG is that only $B = A/\sigma$ is identified, i.e., A is identified up to a scalar. This is mainly irrelevant to study assortativeness on a single market, since comparing different entries of the matrix B is equivalent to comparing the elements of A . Nonetheless, [Ciscato et al. \(forthcoming\)](#) stress that it is not possible to compare the affinity matrices of different markets without a further restriction on A .⁸

Denote A^t the affinity matrix in year t . In order to compare marriage markets over time, we assume that the Frobenius norm of the submatrix \tilde{A}^t is equal to one for every t , i.e., $\|\tilde{A}^t\| = 1 \forall t$. This implies that $\frac{\tilde{B}^t}{\|\tilde{B}^t\|} = \tilde{A}^t$, which in turn implies that $\sigma^t = \frac{1}{\|\tilde{B}^t\|}$. This means that we interpret large global changes in the submatrix \tilde{A} as due to a shift in the relative relevance of unobservables in mating.

Although we need to introduce this further restriction to proceed with cross-market analysis, note that the optimal matching function $\pi(x, y)$ only depends on $B = A/\sigma$. Hence, it stays unchanged under different identification assumptions. This makes the results of our counterfactual analysis of inequality in Section 2.6 robust with respect to different restrictions on the parameters A and σ . We provide a formal proof for this statement in Appendix 2.A.1.

2.3.5. Counterfactual methodology. An interesting, but still unused, feature of DG's model is the possibility to compute counterfactual equilibrium matching by operating on the matrix of preference parameters A . The idea is to infer the marital preferences (A^t, σ^t) from cross-sectional data on couples (X^t, Y^t) for a given year t and then compute the equilibrium matching $P(s, s; t) \equiv \pi(x^s, y^s; A^t, \sigma^t)$ for population data (X^s, Y^s) under the same marital preferences. In this way, by comparing the counterfactual $P(s, s; t)$ to the actual $P(s, s; s)$, we can tell *how people would match if preferences stayed unchanged between period s and t* .

Using $P(s, s; t)$ together with data (X^s, Y^s) , we can compute the counterfactual distribution of couples' characteristics. For instance, we can compute the distribution of household income, as well as various measures of inequality, such as the Gini coefficient. In this way, we can tell to what extent the distribution of household characteristics has changed because of shifts in marital preferences.

⁸Alternatively, one could put an additional restriction on the parameters σ , for instance $\sigma = 1$ on each market. [Ciscato et al. \(forthcoming\)](#) propose to normalise the social gain $\mathcal{W}(A, \sigma)$.

Moreover, it is also possible to create a counterfactual match between subpopulations from different cross-sections. In fact, we can predict the matching $P(s, t; s)$ originating from a fictional situation in which men from year s met women from year t , with the preference parameters s . In this way, it is possible to assess how changes in the marginals influenced the match in order to address specific questions. Although we do not employ this last type of experiment, we recommend it for future research.

While this counterfactual analysis unveils the hidden potential of the model of DG, it also shows an important limitation concerning its empirical application to the marriage market. In absence of a more explicit household model that explains how agents determine their labor supply and advance in their working career, we are forced to consider wage rates and working hours as exogenous characteristics. The counterfactual analysis does not take into account that spouses adjust their labor supply and take on different working careers according to the partners' characteristics and household decision-making process.

2.4. Data

The paper uses CPS data from 1964 to 2017 (March Supplement) from the Integrated Public Use Microdata Series (IPUMS). CPS data provides a detailed representation of the married male and female populations in the United States over time. Hence, they provide us with reliable “photographs” of the marriage market equilibrium we aim to study. In reality, people are likely to meet and marry in small, local marriage markets. Because of the limited sample size of CPS yearly database⁹, we do not account for heterogeneity in sorting patterns across smaller geographical units (such as states or counties) and we present aggregate trends at the United States level.

In this section, we describe the construction of the main variables of interest and the selection of the samples. We also present summary statistics on our population of couples before turning to estimation.

2.4.1. Construction of variables. Our empirical analysis makes use of five key variables: age, education, wage, hours of work and race. In a few cases, such as age, the construction of the variable is straightforward as we take the raw data without further adjustments. In the following, we explain how we deal with other variables.

- *Educational attainment* is available for all years, but with various levels of detail. IPUMS provides a 12-level education variable, which we convert into years of schooling, and to which we refer as the “continuous education variable”. However, this variable is not entirely consistent across years (the coding changed after

⁹In addition, in CPS waves before 1976, there is no state variable at the household level. Only broad geographical areas are reported.

1992). To overcome this difficulty (and provide summary statistics on broader education groups), we constructed two other education variables, one with 5 levels and one with 4 levels.¹⁰ Robustness of the results is checked for each of these specifications.

- As concerns *hours of work per week*, the most consistent variable across waves is “hours worked last week”, as the usual hours of work are not available prior to 1976. However, we check the robustness of our main results obtained with the first definition by implementing checks with the latter, as well as with a combination of the two.
- We define *annual labor income* as the sum of salary, self-employment income and farming income. These components are top-coded. However, as top-coded observations account for only a (very) small fraction of the sample, the results are not affected by the way we deal with them, and we eventually choose to drop them.
- We compute *hourly wages* using labor income, hours of work per week and weeks of work per year.¹¹ We constructed it as follows:

$$\text{wage} = \frac{\text{labor income}}{\text{hours} \times \text{weeks}} \quad (2.4.1)$$

However, the wage variable may feature abnormally low or abnormally high values. We follow [Schmitt \(2003\)](#) advice to trim the data, dropping values below 1\$ or above 100\$ (in 2002 dollars), while keeping observations with a zero wage. All income and wage variables are converted to 1999 dollars.

- There is no consistent *race/ethnicity* variable across years. In the early waves of the CPS data, individuals are only classified as White, Black or Other. After 1971, it becomes possible to separately identify Hispanics and, after 1988, Asians.¹² Across the years, the race variable became more detailed, allowing individuals to declare a mixed ethnic background. However, when comparing preferences across waves, we need to use a consistent specification of the variable. We mainly use three different specifications: (1) Black or White, available since 1962 and considering Hispanics as White after 1971; (2) Black, White, Hispanic and Other, available since 1971 and reallocating Asians into the residual category Others; (3) Black, White, Hispanic and Asian, available since 1988.

¹⁰The 5 levels variable is constructed as follows: (1) below high school degree, (2) high school degree, (3) some college, (4) college degree and (5) 5+ years of college. With 4 levels only, we distinguish: (1) below high school degree, (2) high school degree, (3) college degree and (4) 5+ years of college.

¹¹The number of weeks working in the past year is usually available as a continuous variable. However, it is sometimes only available as a grouped variable, which we use to proxy the number of weeks worked in the past year.

¹²Comparing summary statistics before and after 1971 suggests that most Hispanics declared themselves as White, whereas the category Others mostly contain Asians before 1988.

In most of our specifications, we use five variables, namely age, education, hourly wage, hours of work and race. We test the robustness of our results to the inclusion of other variables (such as occupation) or to alternative coding of the variables.

2.4.2. Sample selection. For every cross-section (i.e., every wave of the survey), we consider the current matches as those resulting from the stable equilibrium of the marriage market. In our empirical analysis of the marriage market equilibrium, we need to decide what matches to include in the sample, which results in several practical issues. First of all, we recall that our analysis of the marriage market equilibrium does not include singles: we exclude never married, separated, divorced and widowed individuals from the sample. In addition, we do not include unmarried couples in our main sample: cohabitation out of wedlock can be a “trial period” before marriage but also an alternative to it, which makes it hard to distinguish the two cases in the data. Couples where spouses live in different households and same-sex couples are also excluded. On the other hand, we do not make any difference between individuals that married once and those who married more than once.

Most importantly, we select couples where at least one of the partners is aged between 23 and 35.¹³ The bracket roughly corresponds to the core of prime adulthood and aims to exclude individuals still at school¹⁴. Although in reality the matches we observe took place at different points in time, we assume that, for each cross-section, individuals aged between 23 and 35 compete on the same marriage market. In this case, marriage markets are not rigidly defined by age brackets: particularly, the age difference between the partners and the age of first marriage may vary greatly. However, our empirical analysis relies on the assumption that sorting dynamics are relatively homogeneous for the age bracket 23-35 for each wave. We also select an alternative subsample of couples where we apply different age cutoffs based on the median age of first marriage in that year (see Figure 2.1): we use this sample to run robustness checks, details are provided in section 2.5.1 and table 2.1.

On this delicate point, we differ substantially from most of the matching literature. For instance, [Chiappori et al. \(2017\)](#) use 2010 Census data to construct the population vectors cohort by cohort. Their method relies on the assumption that each cohort is a separated marriage market.¹⁵ Nonetheless, we aim to estimate the intensity of assortativeness on age and document its trend over time. The selection criterion proposed by [Chiappori et al. \(2017\)](#), instead, assumes an extremely rigid sorting pattern with respect to age. This is

¹³Similar simple selection criteria by age are common in the literature. See [Schwartz and Mare \(2005\)](#) (where the wife must be between 18 and 40) or [Schwartz and Graf \(2009\)](#) (where both partners must be between 20 and 34).

¹⁴We also exclude students aged more than 23 by combining data on school attendance and reasons for not participating to the labor market.

¹⁵More precisely, each cohort t of young men matches with the cohort $t + 1$ of young women.

a well-known limitation in the matching literature: including age among the matching variables is a first attempt to deal with this problem.

One of the main concerns affecting age restrictions is the self-selection due to divorce. Separation and divorce allow to observe only the prevailing unions at a given point in time and this may lead to some problems in the interpretation of the results. For example, cohorts born in 1950 have been largely affected by changes in divorce laws in the 1970s, and their divorce rate is particularly high. Divorce may primarily destroy non-assortative matches. Hence, the marriage patterns observed in 2010 for this cohort might result from a selection process through divorce, instead of being the result of the specific tastes at the moment of the match. In order to overcome this potential bias, it could be advised to work with a subsample of newlyweds (as also suggested by [Schwartz and Mare, 2005](#)). Unfortunately, in most cases, data on marital history are not available.

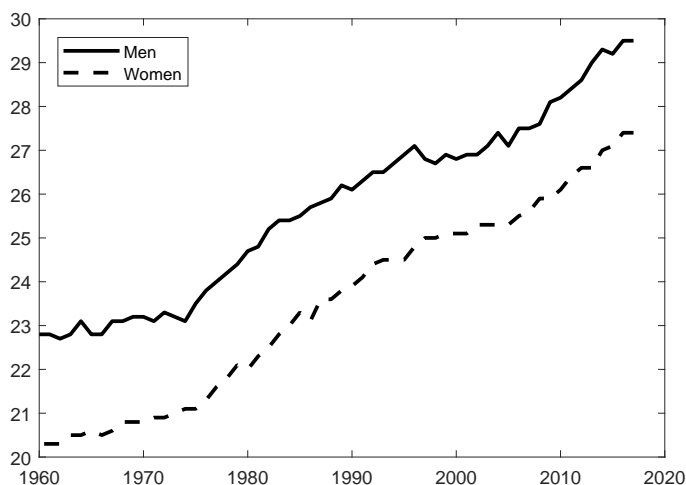
Finally, note that the estimation algorithm works best with samples with order of magnitude equal to 3. For some waves, the sample of observations respecting our selection criteria is greater than 10,000. In Appendix 2.A.2, we propose a methodology to ensure that the sample is highly representative of the sorting patterns when we must reduce its size.

2.4.3. Baseline sample. The changes in the availability of data and potential problems arising from the construction of the variables motivate the use of alternative samples. In spite of this, we choose three baseline specifications described in Table 2.1 in Appendix 2.B that we use to present our main findings. Sample A covers all waves from 1964 to 2017, but employs a limited specification of the race variable: only Blacks and Whites are distinguished. Sample B covers waves from 1971 to 2017, but contains a more detailed race/ethnicity variable: the available categories are Blacks, Hispanics, Whites, and Others. Sample C only covers waves from 1988 to 2017, but, with respect to Sample B, includes an additional racial category for Asians. We introduce three different baseline samples since data about race and ethnicity are not consistent across years. We are primarily concerned with potential biases due to the misspecification of ethnic traits and the exclusion of minorities from the sample. While discussing our main findings, we run several robustness checks that employ different subsamples. These are also described in Table 2.1 in Appendix 2.B.

2.4.4. Summary statistics. The population that we consider in the empirical analysis has gone through major changes in the past fifty years. The median age at first marriage has increased over this period, from 22.8 years old (resp. 20.3 years old) for men (resp. women) in 1960 to 29.5 years old (resp. 27.4) in 2017. Trends for the median age

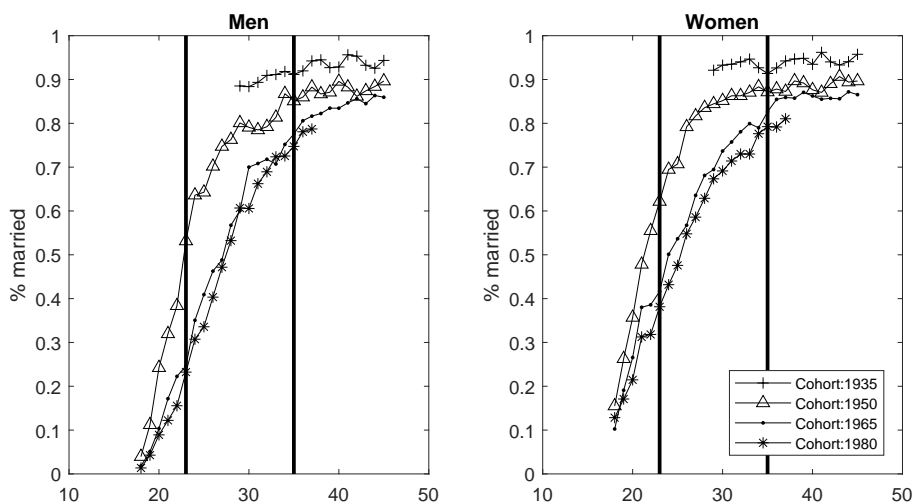
of first marriage are displayed in figure 2.1¹⁶. At the same time, the share of unmarried individuals has greatly increased (see Figure 2.2).

Figure 2.1. Median age at first marriage



Notes: the trends for the median age at first marriage have been estimated with CPS data by the Fertility and Family Statistics Branch of the US Census Bureau and are available online: <https://www.census.gov/data/tables/time-series/demo/families/marital.html>.

Figure 2.2. Share of married individuals by age and cohort



Notes: share of married couples by age and cohort using CPS data 1964-2017. For an individual to be in the data, he must be aged between 18-65. Couples where one partner is still at school are excluded. The vertical lines represent the age interval (23-35) used for selecting our main samples.

The rise in educational achievement is depicted in panel (a) and (b) of Figure 2.3. Only a relatively smaller fraction of individuals now belongs to low education categories (below

¹⁶The trends for the median age at first marriage have been estimated with CPS data by the Fertility and Family Statistics Branch of the US Census Bureau and are available online: <https://www.census.gov/data/tables/time-series/demo/families/marital.html>. Details about the estimation method can be found in Siegel and Swanson (2004, Chapter 9). The 1970 and the 1980 Census and the ACS contain survey data about the year of marriage: a comparison of the estimated CPS trends with these data sources reveals that the estimates are indeed accurate.

high school or high school degree), while an increasing share of the population falls into higher education categories (some college, college degree or above). Note that this trend is especially striking for women, who now appear to be more educated than men, while the reverse was true in the 1960s.¹⁷

In panel (c) and (d), we describe the racial composition of our sample. We can separately identify the four major racial groups (White, Black, Hispanic and Asian) after 1988. From the graph, it seems that Hispanics used to declare themselves as White prior to 1971, while Asians composed the majority of the “Others” category. The share of Black in the samples is relatively constant, while Hispanics and Asians account for an increasing share of the population at the expense of the White category.

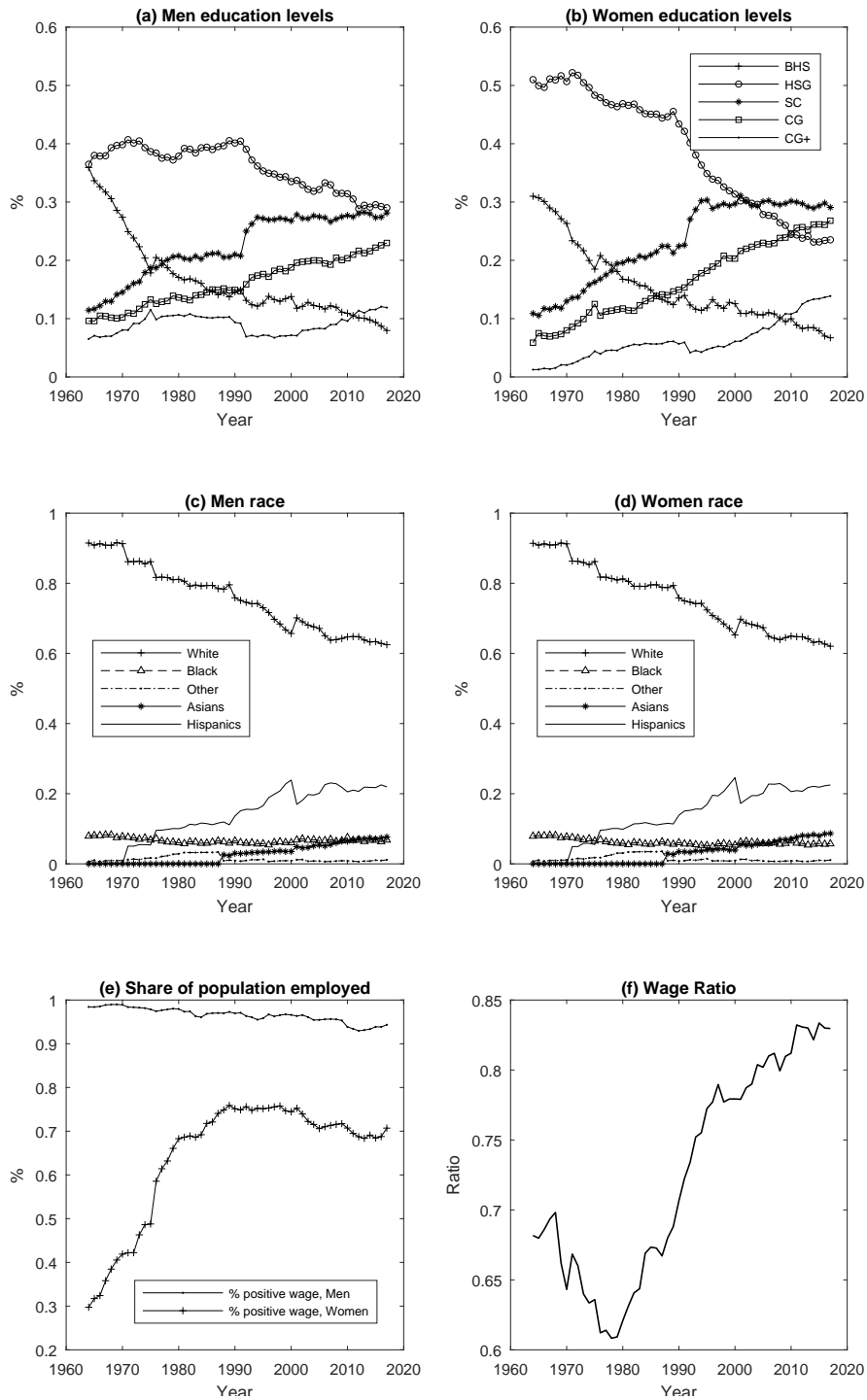
One major change in families in the past fifty years is the increased participation of women on the labor market. This is represented in panel (e) of Figure 2.3. Our measure of employment for our sample is the share of people with a strictly positive wage¹⁸: the graph shows a dramatic increase for women, although the rate stabilised after 1990. Finally panel (f) depicts the wage ratio for women relative to men (conditionally on having a strictly positive wage). This increase has been identified as one of the main factors of change for families (see [Becker, 1973, 1991](#), on specialization within households and human capital investment of women).

When we look at the joint characteristics of the spouses (Figure 2.4), we notice a strong positive correlation between the partners’ age and education, which is a first hint that these traits are complements. While correlation by age decreases over time, the trend of correlation by education is instead unclear. On the other hand, we observe an increasing trend for the correlations by hours worked and hourly wage. Interestingly, these comoments are first weakly negative and then weakly positive. Finally, the share of interracial marriages has increased over time.

¹⁷The graph also shows the discrepancy in the education variable in 1992, as a large share of the population previously categorised as having a high school degree now appears in the “Some College” category.

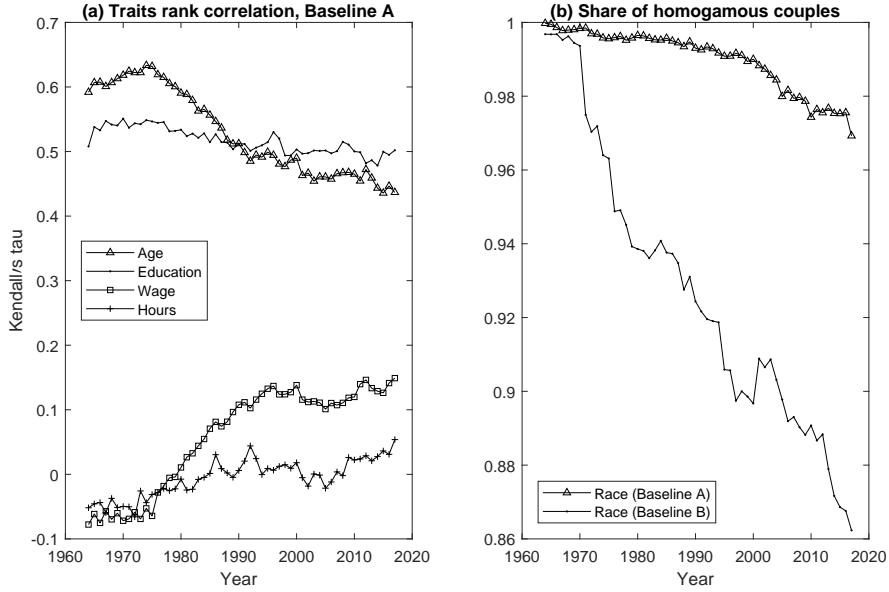
¹⁸The share of employed people may appear extremely high in some cases (for men at the beginning of the period for example), but this may be due to our sample selection criteria based on age and marital status. In addition, we consider a person as employed as long as we are able to compute a wage, that is, as long as she worked in the past year.

Figure 2.3. Summary Statistics



Notes: married couples from CPS data 1964-2017. For a couple to be in our sample, at least one partner must be aged between 23-35. Couples where one partner is still at school are also excluded. Discontinuity around 1992 for schooling trends is due to a change in the variable specification made by the US Census Bureau. Discontinuities in the race trends are also due to the addition of new categories in the set of possible answers. We use 5 levels of education: (BHS) below high school degree, (HS) high school degree, (SC) some college, (CG) college degree and (CG+) 5+ years of college.

Figure 2.4. Partners' Traits



Notes: samples used: baseline A. Baseline sample B is used for the trend of interracial marriage (right panel).

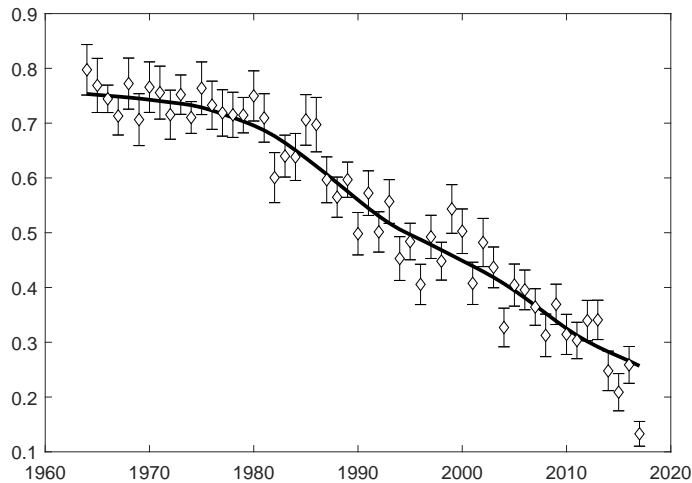
2.5. Trends in matching patterns

In this section, we describe trends for the diagonal elements of the affinity matrix estimated using the baseline sample A described in Section 2.4.3 over the period 1964-2017. Estimation follows the steps explained in Section 2.3. Estimates of the A_{ij} entries are obtained for every year and shown in the graphs below. We display the point estimates, as well as the confidence intervals. Data are standardized so that the covariance matrices $E_{\hat{\pi}}[x'x]$ and $E_{\hat{\pi}}[y'y]$ have diagonal entries equal to one for a reference year¹⁹: this allows us to compare different estimates of A within and across years. We also use local constant regression smoothing (LOWESS) to ease the interpretation of the results. Finally, we present several robustness checks in order to understand whether our baseline findings suffer from variable misspecification, sample selection or endogeneity problems: we provide an exhaustive list of the checks in Appendix 2.B.

2.5.1. Age. Our results show that spouses' ages are strongly complementary. However, Figure 2.5 also shows an unambiguous decrease in age assortativeness. This may appear as in contrast with previous results by [Atkinson and Glass \(1985\)](#) and [Qian and Preston \(1993\)](#), who claim that in the United States homogamy by age increased up to 1987. Nonetheless, this trend could be explained by a progressive passage from a traditional form of marriage - where the woman is slightly younger than the man - to a variety of different unions. For instance, [Atkinson and Glass \(1985\)](#) notice that

¹⁹In practice, we first compute $(diag(E_{\hat{\pi}}[x'_{1991}x_{1991}]))^{-1/2}$ for men's population in 1991, and then use it as a scaling factor for every cross-section. Same for women's population.

Figure 2.5. Assortativeness in age



Notes: sample used: baseline A. The figure displays the estimated trend of the diagonal element of the marital preference parameter matrix A capturing the interaction between husband's and wife's age. We observe a decrease in age complementarity.

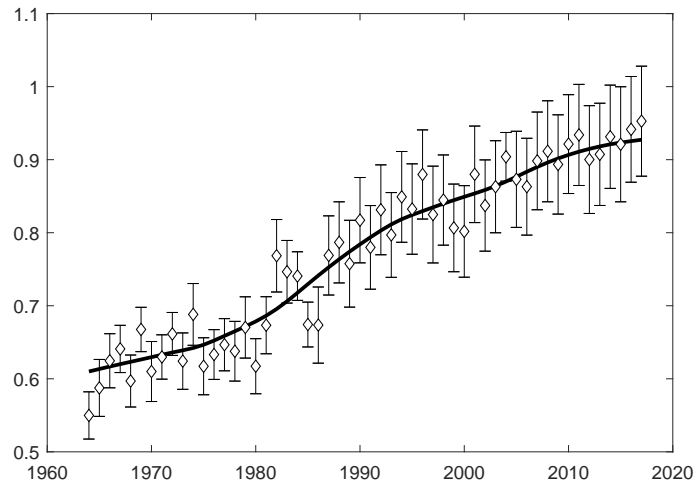
spouses with similar socio-economic background tend to be of the same age more and more frequently. Moreover, couples where the husband is younger or where the difference in partners' ages is high are more and more socially acceptable. What we find is, in fact, that the *strength of sorting* decreased, which means that several age combinations now coexist at equilibrium.

The age bracket we consider is large enough to include most couples married in the years preceding the survey year we look at.²⁰ However, since our age cutoffs are fixed (the youngest spouse must be at most 35), and since people tend to marry later and later over the time period considered, the composition of our sample in terms of marriage duration is likely to change over time. To address this issue, we consider an alternative sample made of couples where at least one partner's age is in the interval $[\bar{a}_t - 4, \bar{a}_t + 2]$, where \bar{a}_t is the median age of first marriage for men in year t .²¹ Sorting trends obtained with this sample present some differences with respect to those obtained with our baseline sample (see Figure 2.16). These differences are most likely due to the different composition in terms of year of birth and marriage duration. However, note that all qualitative results hold (the slope of the trends is the same as in our main findings). In particular, the strength of sorting with respect to age is lower, since the sample's age range is reduced, but the trend is still decreasing.

²⁰In fact, in spite of the increasing median age of first marriage, most people still get married before they turn 35. The share of married people in a cohort is indeed approximately constant after 35 (see Chapter 1, [Browning et al., 2014](#)).

²¹The difference between the median age for men and for women is approximately constant over the time period considered, and is around 2 years (see Figure 2.1).

Figure 2.6. Assortativeness in education

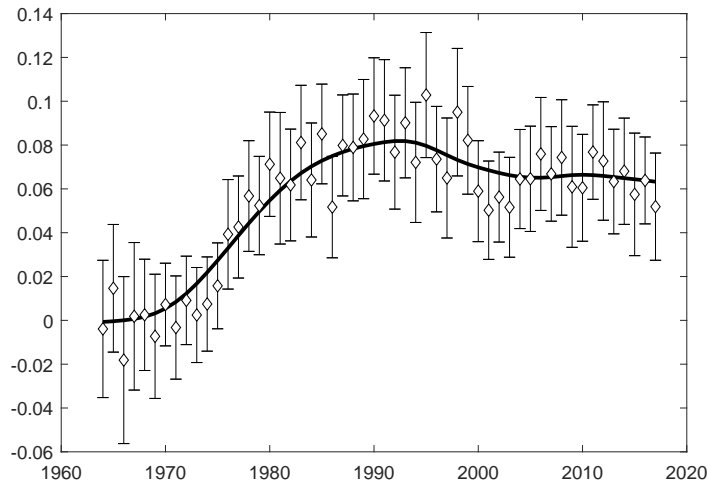


Notes: sample used: baseline A. The figure displays the estimated trend of the diagonal element of the marital preference parameter matrix A capturing the interaction between husband’s and wife’s schooling level. We observe an increase in education complementarity.

2.5.2. Education. Figure 2.6 represents the trend of assortativeness in education between 1962 and 2017. We find a general increase in assortativeness in education. This is in line with the results of [Greenwood et al. \(2014\)](#), and with most of the findings in the literature (see Section 2.2). Nonetheless, as explained throughout the paper, we argue that our estimates only capture marital preferences and are cleansed from any demographic effect. These findings also provide further support to those of [Chiappori et al. \(2017\)](#), who document a rise in educational assortativeness for cohorts born between 1943 and 1972: while they also take a structural approach, they need to assume a fixed sorting structure for age and are limited to a unidimensional matching. Since we work in a multivariate setting, we can “control” for other observables and also conclude that assortativeness in education is comparable in strength to age, whereas it is much higher than wage or hours of work (see Figures 2.7 and 2.8). As concerns possible misspecification of the schooling level variable, we find that our results are robust to different measures of educational attainment (detailed findings available on request).

2.5.3. Wage. The estimates for wage assortativeness are presented in Figure 2.7. In the earliest waves, the estimates of the affinity matrix parameter for wages are not significantly different from zero. However, assortativeness in hourly wage rates has steadily increased up to the 2000s and is significantly positive in every wave since the mid-1980s. In the last 15 years, it seems that the estimate stabilised around a value of 0.05. The trend for the wage estimate is parallel to the closing of the wage gap and may suggest that men developed a stronger incentive to look for a spouse among high earners. Since our estimation includes both hours of work and hourly wages among the matching variables, and since we construct hourly wages using hours of work, we test the robustness of our

Figure 2.7. Assortativeness in wage



Notes: sample used: baseline A. The figure displays the estimated trend of the diagonal element of the marital preference parameter matrix A capturing the interaction between husband's and wife's wage. We observe a possible rising of a relatively weak wage complementarity which was not observed in the early waves.

results using annual earnings instead of wages.²² The results are displayed in figure 2.17 in the appendix and are very similar to those obtained with our baseline sample.

In spite of this (weak) positive assortativeness, [Becker \(1973\)](#) suggested that the spouses' wages should be substitutes because of household specialization, while main non-labor-market traits are expected to be complements. Unfortunately, our result is not a good test for Becker's predictions: since many women (as shown in panel (e), Figure 2.3) are not part of the labor force - especially in the earlier waves - we are not able to observe their wage potential. In other words, we are not able to determine the *shadow price* of time spent away from the labor market to which Becker refers to in his analysis of the household. As a result, the estimates we present do not capture marital preferences because of this endogeneity issue affecting the observed hourly wage rates (see also [Ciscato et al., forthcoming](#)).

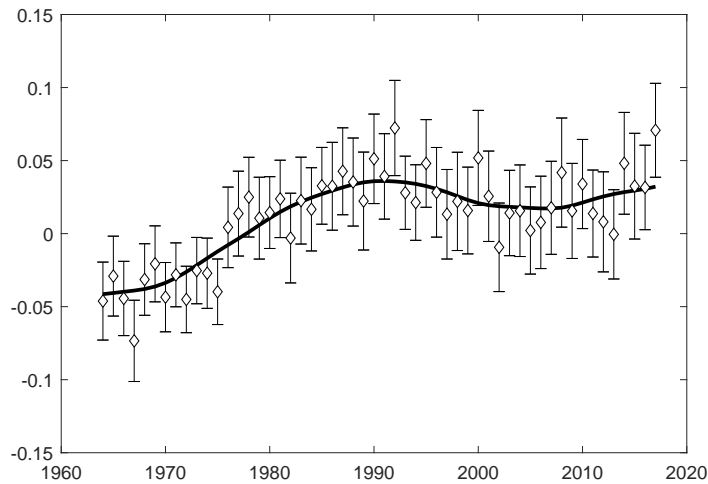
To understand to what extent our main findings are affected by endogenous workforce participation choices, we run two parallel estimations with different subsamples (see Appendix 2.B, checks 3 and 4). First, we only estimate the affinity matrix for a subsample of couples where both spouses have a positive wage. With respect to our baseline results (Figure 2.7), we find evidence of positive assortative mating on wages since the earliest waves and the strength of assortativeness is now constantly larger (Figure 2.18). Second, when we only consider the subsample of childless couples, the estimates for wages' complementarity are even higher (Figure 2.19). These checks seem to suggest that, for households where household specialization is expected to be less pronounced,

²²Precisely, we replace wages by the variable $\log(1 + w^a)$ where w^a are (possibly 0) annual earnings.

sorting on spouses' wages is indeed stronger. Nonetheless, as both fertility and labor force participation are the outcomes of endogenous choices, none of the two subsamples can be considered as representative of the population preferences. Further research is needed in this direction.

Finally, we construct a measure of potential income that allows us to deal with non-participation (see Appendix 2.B, check 5). We predict wages using a standard selection model (Heckman, 1979). We estimate the trends of marital preferences and report the result in Figure 2.20. The main important finding is that we observe positive assortative mating on wages since the earliest waves. The trend is increasing with a slope comparable to the one obtained with baseline sample A. We discuss the potential implications on inequality trends in Section 2.6.5 in light of the results of our decomposition exercise.

Figure 2.8. Assortativeness in Hours of work



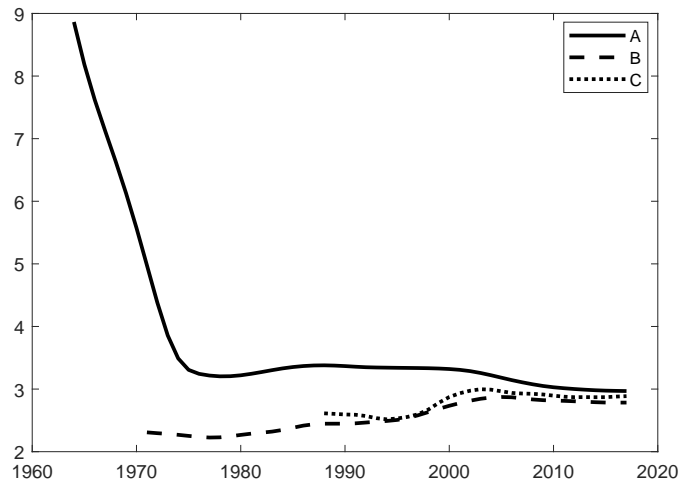
Notes: sample used: baseline A. The figure displays the estimated trend of the diagonal element of the marital preference parameter matrix A capturing the interaction between husband's and wife's hours worked. We observe a possible rising of a relatively weak complementarity in hours worked which was not observed in the early waves.

2.5.4. Hours Worked. Trends in mating preferences for hours of work are represented in Figure 2.8. Similarly to the case of wages, the estimates for the earliest waves are not significantly different than zero and are even negative for some waves, whereas we observe an irregular increase starting from the 1980s. The increasing trend seems consistent with the shift from production complementarities as the main source of marriage gains to consumption-based complementarities (Stevenson and Wolfers, 2007): while in traditional families one spouse - typically the wife - focused on housework and the other on the labor market, now partners may benefit from similar time schedules.

Once again, what the estimate for hours worked captures cannot be interpreted in terms of preferences at the moment of the match, since spouses most likely adjust their labor supply after the marriage. Checks 3 and 4, described in Section 2.5.3, lead to the following results:

for couples where both partners are employed, we observe positive assortative mating on time schedules for any wave (Figure 2.18), while for childless couples the positive sorting is even stronger (Figure 2.19). In both cases, the strength of complementarity increases over time, similarly to the baseline trend. Although these estimates are biased because of selection, it seems that couples where household specialization is weaker indeed display more homogeneous working time schedules and leisure time spent together.

Figure 2.9. Assortativeness in Race, by number of race included



Notes: samples used: baseline A, B and C (see Appendix 2.B). The figure displays the estimated trend of the homogamy preference parameter for race contained in the matrix A. We observe an increase in the preference for racial homogamy.

2.5.5. Race. Figure 2.9 reports our estimates of the racial homogamy parameter for baseline sample A, B and C, described in Section 2.4.3. We observe a sharp decline in the taste for homogamy when considering the race specification Black-White: the most significant decrease took place during the 1960s, when the last anti-miscegenation laws were ruled unconstitutional, whereas we observe a steady but only slight decrease from the 1970s. Interestingly, when switching to the specification Black-White-Hispanic-Others (sample B), the trend is instead slightly increasing over the period 1971-2017. Finally, when the specification Black-White-Hispanic-Asian-Others is considered (sample C), the trend is increasing up to the mid 2000s, and then slightly decreasing. In general, however, the results depend on how many groups are considered, that is, on the level of detail of the classification scheme. Studies on racial homogamy are facing the same issue, as the number of racial groups may vary depending on the availability of the data, or on how individuals are allowed to report their race. We can conclude that, although the data show a growing number of interracial marriages, the latter became less desirable since the 1970s when considering a detailed level of ethnic fragmentation, and that only recently the trend might have been reversed.

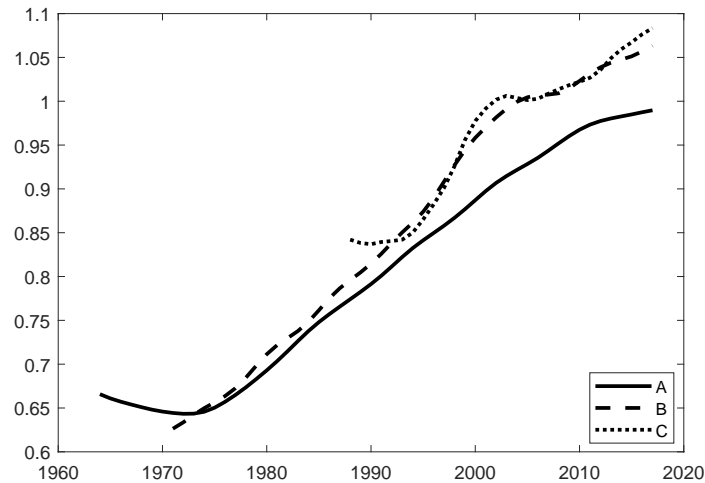
We can further read into our results by switching from our main parametric specification (2.3.5) to an alternative one, where, instead of treating race as a categorical variable denoted $x_R, y_R \in \{White, Black, Hispanic, Other\}$, we include one dummy for each of the four racial groups. In other words, given our 4 racial categories, we restrict to zero the diagonal elements representing the interaction terms of type, say, $\mathbb{1}\{x_R = White\}\mathbb{1}\{y_R = White\}$ and are left with $4 \times 4 - 4$ off-diagonal parameters to estimate. In this way, in our robustness check 6, we are able to identify the surplus change resulting from marrying a partner with a different racial background relatively to matching within one's own racial group. Trends are indeed heterogeneous by racial group (see Figure 2.21). Blacks appear as the most segregated racial group, suffering the highest penalties for interracial marriage on the market. However, most trends associated with interracial couples where one partner is Black suggest that these interracial marriages have grown more attractive (or at least not less attractive). On the other hand, Hispanics and Whites seem to have grown apart on the marriage market, with the trend only being reversed from late 2000s. Hence, this should explain our original conclusion: interracial marriage is found to be less and less attractive over time when including Hispanics as a separate ethnic group.

Our findings are particularly counterintuitive because the share of interracial couples has been growing as the American population becomes more and more multi-ethnic. In this case, disentangling preferences from demographics leads us to a result in contrast with conclusions from previous works. [Fryer \(2007\)](#) uses a specification White-Black-Asian for his race variable and concludes that preference for homogamy decreased throughout the last century: this seems to be in line with our estimates obtained with baseline sample A, where Hispanics are mainly considered as Whites. However, [Fu and Heaton \(2008\)](#) uses a White-Black-Hispanic-Asian specification to describe ethnic groups: this time, our results are opposed to theirs, as they conclude that taste for racial homogamy decreased over the period 1980-2000.

2.5.6. Unobservables. We recall that the scalar σ measures the relevance of the unobservable random preference shock: the higher σ , the more matching appears as random to the observers. Figure 2.10 displays the values of σ obtained under our identification assumption given in Section 2.3.4 and with baseline sample A. The clear increasing trend suggests that socio-economic observables matter less today than they did fifty years ago. The role played by the parameter σ in our theoretical framework suggests that there are two forces offsetting each other. On the one hand, we report that taste for racial homogamy and positive assortativeness in education have increased in strength. On the other hand, the relevance of the socio-economic observables that we take into account has decreased.

In robustness check 7, we purposefully omit education in order to understand how this would bias our findings, and particularly how this would change the relative importance

Figure 2.10. Sigma



Notes: samples used: baseline A, B and C (see Appendix 2.B). The figure displays the estimated trend of the parameter σ capturing the relevance of idiosyncratic preference shocks in our matching model. We observe an increase of the relevance of unobservables in matching.

of the entropy observed by the econometrician as measured by σ . This exercise shows that omitting a key variable such as education results in upward-biased estimates of complementarities of traits that are positively correlated with the omitted variable. In this case, omitting education results in higher estimates for complementarity on age and wage, as well as for preference for ethnic homogamy (see Figure 2.22). In addition, σ is found to be higher, and its trend increases more steeply: in this omitted-variable specification, unobservables play a more relevant role, and their importance grows faster than in our baseline specification.

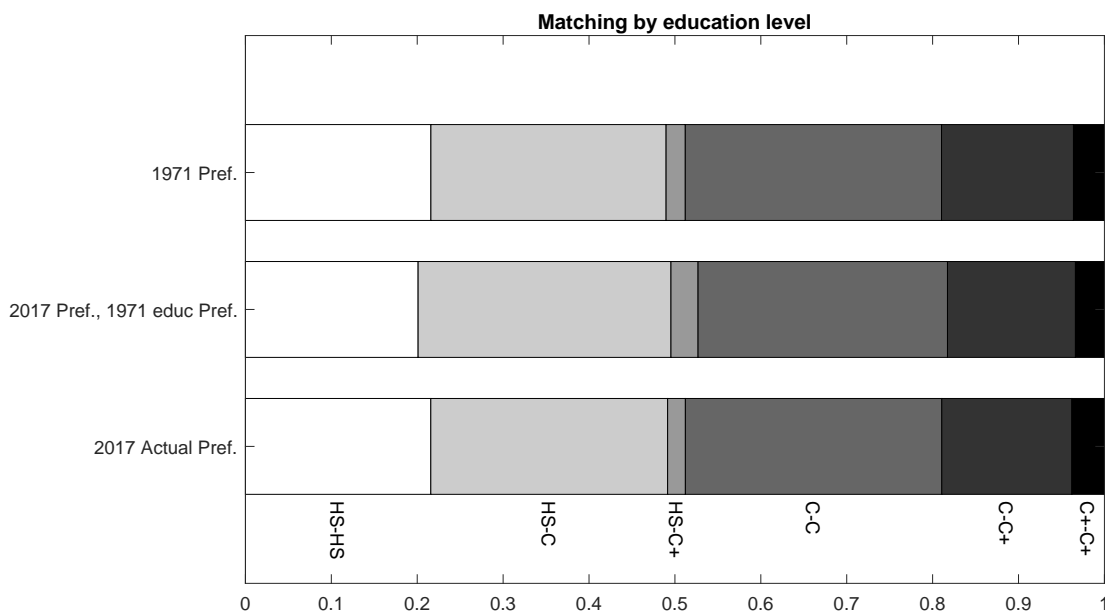
2.6. Counterfactual analysis

One key motivation behind the analysis of marital preferences is to understand their contribution to the changes in mating patterns and between-household inequality. To conduct our counterfactual analysis, we used CPS data for the years 1971 and 2017. We use the baseline sample B, which includes four racial groups (White, Black, Hispanic, Others). We try to answer two questions: (a) what would be the marriage patterns, for example the joint distribution of education, if individuals married as in 1971? (b) how inequality would change if individuals had the same marital preferences as in 1971?

As explained in Section 2.3.5, once estimated A^{1971} and with (X^{2017}, Y^{2017}) at hand, we can compute the counterfactual optimal matching $P(2017, 2017; 1971)$ in order to compare it with the marriage market equilibrium predicted by the model with the actual 2017

preferences²³. In this section, we report the results of our counterfactual analysis for two variables, age and education. We subsequently proceed with the analysis of inequality trends and a decomposition exercise to understand which parameters are associated with the inequality rise.

Figure 2.11. Assortativeness in education, counterfactual distribution



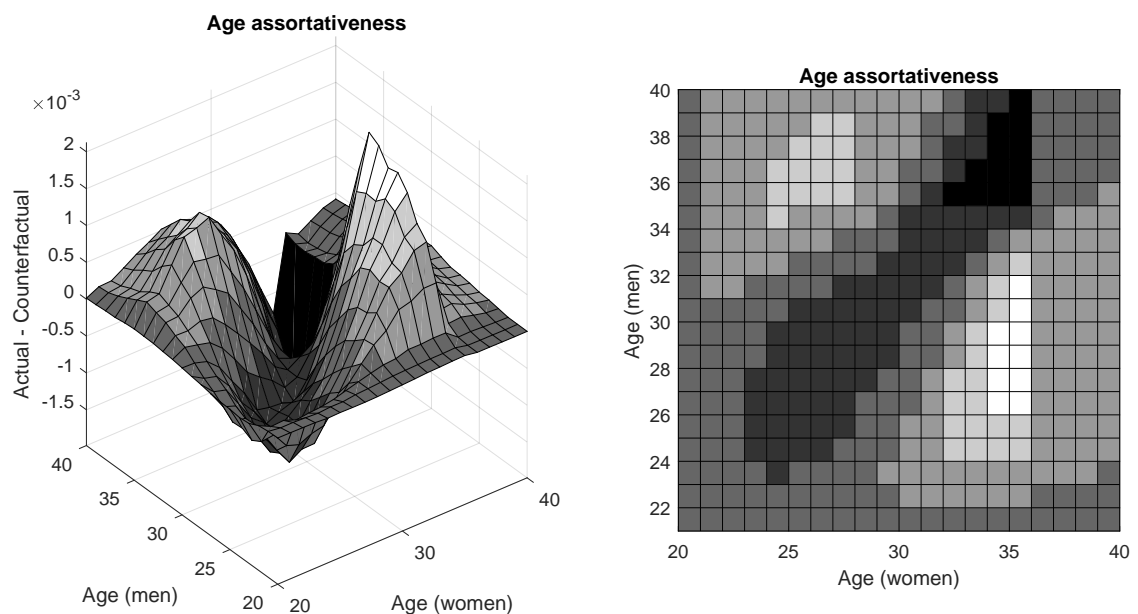
Notes: sample used: baseline B. The marginal distributions of characteristics (X^{2017}, Y^{2017}) are taken from 2017 data for the three figures. In the first line, we show the counterfactual joint distribution of partners' educational levels (high school and below (HS), some college (C), and college degree and above (C+)) obtained using 1971 marital preferences. In the second line, we show the counterfactual distribution obtained using 2017 marital preferences but allowing the schooling complementarity parameter to be equal to its 1971 value as in Figure 2.6. In the third line, we show the actual distribution obtained with 2017 marital preferences.

2.6.1. Education. To ease the representation of the results, we gather individuals in three educational types: high school and below (HS), some college (C), and college degree and above (C+). We first compute the counterfactual market equilibrium $P(2017, 2017; 1971)$: the upper distribution in Figure 2.11 displays the relative frequencies of each of the six possible types of match that would result from matching if the preferences of the 2017 population were the same as in 1971. The relative frequencies reported in the second line are the result of a different counterfactual experiment. We now fix all the parameters to their 2017 values except for the one capturing the interaction between partners' education, which is allowed to take its 1971 value. In this way, we isolate

²³Alternatively, we could compare the counterfactual matching $P(2017, 2017; 1971)$ with the actual frequencies observed in the data, rather than those predicted by the model. However, since the model does an excellent job in reproducing the actual frequencies, the two exercises yield similar results.

the effect of the change in this single parameter on the marriage market outcome. In the last line, we report the joint aggregate equilibrium distribution of educational types predicted by the model for year 2017. Comparing line 2 and 3, we note that the increase in complementarity shrinks the shares of couples crossing educational barriers. On the other hand, endogamous marriages are more frequent. With the help of line 1, we conclude that these changes observed above turn out to be mostly offset by shifts in other parameters, which leads us to conclude that the evolution of marital preferences had little impact on the joint distribution of partners' schooling levels. One main reason is likely to be the increase in the parameter σ , which decreases the relevance of socio-economic observables on the marriage market.

Figure 2.12. Assortativeness in age, counterfactual distribution



Notes: sample used: baseline B. The figures depict the differences in joint frequencies of partners' ages between the actual distribution obtained with 2017 preferences and the counterfactual one obtained with 1971 preferences ($P(2017, 2017; 2017) - P(2017, 2017; 1971)$). We show such frequencies in a three-dimensional space and in the corresponding “elevation” map. The darker the block, the more couples of the corresponding age in the counterfactual outcome outnumber their peers in the actual. The lighter the block, the more couples of the corresponding age in the actual outcome outnumber their peers in the counterfactual. Remember that the sample may include individuals of any age, although we require that at least one partner is aged between 23 and 35 for the couple to be in the sample.

2.6.2. Age. We repeat a similar experiment with age, as illustrated in Figure 2.12, where we computed the joint distribution of spouses' age with both actual and counterfactual marital preferences, then we looked at the difference between the two. Remember that, as discussed in Section 2.5.1, in 1971 there used to be a relatively stronger sorting on

age than in 2017. From Figure 2.12, we note that, under counterfactual 1971 preferences, we would observe far more couples where the husband is around 2 or 3 years older than the wife (the darkest cells are mostly right above the diagonal) and slightly more couples with partners of the same age (the cells on the diagonal are dark). These two types of couples, and especially those with a slightly older husband, can be considered as the most “traditional”. However, the change in preferences made them less frequent in favor of other types of marriages. Indeed, under actual 2017 preferences, we observe far more couples where the distance between spouses’ ages is greater (the white cells are far from the diagonal). In particular, there are many more couples where the wife is more than 5 years older than the husband.

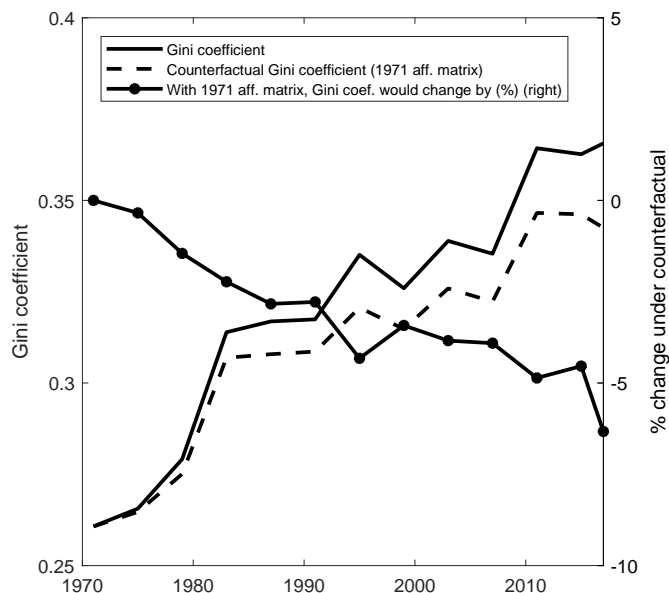
2.6.3. Inequality. The purpose of the previous sections was to show how the change of the affinity matrix directly translates into a different marriage market outcome. We can now compute household income distributions and then Gini coefficients. For each potential couple, we compute the total labor income of the household, while the optimal matching matrix $P(s, s; t)$ tells us the corresponding frequency of this type of couple at equilibrium. For any two years s and t , we use individual traits distribution from year s and marital preference parameters from year t and compute the Gini coefficient using the optimal matching matrix $P(s, s; t)$ - i.e., the counterfactual frequency table of the couples’ type - and the vectors of spouses’ incomes x^s and y^s . We denote $\mathcal{G}(s, s; t)$ the predicted Gini coefficient computed with male and female population vectors from year s and with marital preferences (A^t, σ^t) from year t .

We aim to study the evolution of inequality from 1971 to 2017. In particular, we ask the following question: *what inequality patterns would we observe in year s if the year s population had the same marital preferences as in 1971?* To answer, we first compute the Gini coefficient predicted by the model for every year s using preferences in year s , $\mathcal{G}(s, s)$. Then, we fix marital preferences to their 1971 levels but predict marriage patterns for the population in year s . Hence, we can compute $\mathcal{G}(s, 1971)$ using the counterfactual labor income distribution. We replicate this exercise every four years starting from our reference year 1971 and always fixing marital preference parameters to their 1971 levels. In figure 2.13, we plot the predicted Gini coefficients with actual preferences (solid line) and with counterfactual preferences (dashed line), while the dotted line depicts how the Gini coefficient would change (in percentage) if individuals had the same tastes as in 1971. While inequality is steadily increasing since 1971, the solid and dashed lines slowly diverge from each other, which means that the rise of household income inequality has been exacerbated by the shifts in marital preferences.

Similarly to [Eika et al. \(forthcoming\)](#), we observe a clear increase in income inequality among married households over the last 45 years, from 26.76 points in 1971 to 36.56 in 2017. However, were marital preferences constant since 1971, the current Gini would be

lower by 2.43 points (equal to 34.13, that is about 6% less). Our experiment indicates that 24.80% of the rise in inequality in total household labor income between 1971 and 2017 can be attributed to changes in preferences on the marriage market.

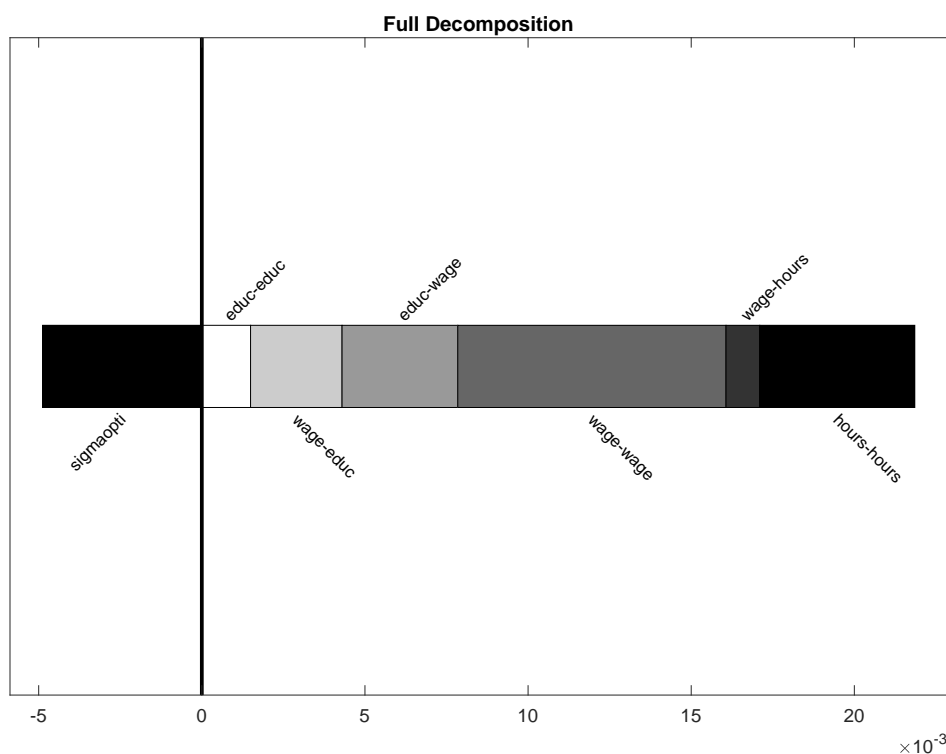
Figure 2.13. Counterfactual analysis, Gini coefficients since 1971



Notes: sample used: baseline B. The figure shows the estimated actual trend of the between-household Gini coefficient and a counterfactual trend obtained by fixing marital preferences to their 1971 values.

2.6.4. Decomposition. Finally, we decompose the share of the increase of the Gini coefficient that we attribute to shifts in marital preference parameters (Figure 2.14). On the right of the vertical axis, we find the main forces that contributed to the rise of household inequality. Not surprisingly, increased complementarity in partners' education is one of them, albeit not the strongest. In fact, despite the modest size of their increases (see Figures 2.7 and 2.8), the changes that have concerned sorting on wage rates and working hours seem to be the main driving forces behind the inequality rise. However, even small variations in the parameters may result in large fluctuations of macroeconomic outcomes if the marginal distributions change. Since the wage structures, the wage gender gap and women's participation to the workforce have radically changed (see panels (e) and (f) in Figure 2.3), the interaction of such transformations with marital preference evolution has amplified inequality growth. In addition, an important share of the change in Gini coefficient is due to shifts in cross-interactions, i.e., of those parameters that do not lie on the diagonal of the affinity matrix. In particular, the interaction between one's education and his/her partner's wage and the interaction between husband's wage and wife's hours worked play a prominent role. Some of these trends can be found in Appendix 2.C, Figure 2.15, and once again our estimates suggest that such interactions are relatively

Figure 2.14. Decomposition of Gini coefficient shift 1971-2017 due to marital preferences



Notes: sample used: baseline B. In the labels, the first trait is the husband's and the second is the wife's, e.g. Wage-Educ refers to the interaction between the husband's wage rate and the wife's education. On the right of the vertical axis, there can be found the parameters that contributed to raise inequality; on the left of the vertical axis, those that pushed in the opposite direction, leading to a decrease.

weak and have not changed much over time²⁴. Finally, looking at the left of the vertical axis, we find that the increase of the parameter σ has hampered inequality by reducing the relevance of socio-economic observables in matching. Another counterforce to the rise of inequality is the decrease in assortative mating on age²⁵.

2.6.5. Discussion. In our counterfactual analysis, we consider changes in sorting along several dimensions, not only education, and show that changes in marital preferences must be regarded as an important driving force behind the recent rise in inequality. We conclude that changes in sorting account for almost 25% of the total increase in household income inequality among married couples, as measured by the Gini coefficient. In line with theoretical predictions of [Fernández et al. \(2005\)](#) and previous empirical findings, educational assortativeness is shown to have a positive impact on income inequality.

²⁴The only trend that we find is a slight - and barely significant - decrease of the negative interaction between men's wage and women's working hours. This parameter may partly capture the wife's wealth effect on labor supply.

²⁵However, the effect is small and is not displayed in figure 2.14.

Greenwood et al. (2014) and Eika et al. (forthcoming) find that changes in educational assortativeness had almost no direct impact on inequality, and point at changes in labor market participation and returns to education as the main forces explaining household income inequality trends. While we do find a positive effect of increasing educational sorting on inequality, we also conclude that this effect is of second order when other controls are included among the matching variables.

Greenwood et al. (2016) suggest that, all else constant, changes in the marginal distribution of hourly wages are key in understanding the increase in household income inequality. Our findings suggest that changes in the marginal distribution of wages paired with a small, but significant, increase in the strength of sorting along hourly wages and hours worked account for a sizable share of the inequality rise. While, unlike Greenwood et al. (2016), we are unable to discuss the role of endogenous labor-supply adjustments, we run several robustness checks to understand how the endogeneity of labor supply choices can bias our findings. Results obtained with a sample of two-earner households (Figure 2.18), childless couples (Figure 2.19) and with potential income as a matching variable (Figure 2.20) lead us to the following considerations: first, our main estimates on complementarities on hourly wages and hours worked are likely to be downward biased; second, the same robustness checks confirm that the direction and the size of the changes are mainly unaffected by the bias.

Finally, Greenwood et al. (2014) and Greenwood et al. (2016) suggest that the uneven decline of marriage has further contributed to the rise of inequality across households, as individuals with low education have become comparatively less likely to be married in the cross-section. In this paper, we exclude the extensive margin from the analysis - i.e., the choice of marriage vs singlehood - and thus we do not relate changes in assortative mating to the decline of marriage, consistently with the logit framework of DG.²⁶ In a related project, Dupuy and Weber (2018) discuss the importance of the extensive margin as opposed to the intensive margin in a model of educational assortative mating: while changes along the extensive margin play a greater role, changes along both margins positively contribute to the rise of inequality.

2.7. Conclusions and perspectives

Our analysis calls into question and updates previous results on the evolution of mating patterns and their implications for inequality: it aims to provide the most recent and complete picture of mating patterns in the United States relying on a structural approach

²⁶As discussed by DG (see Appendix D in particular) and Ciscato et al. (forthcoming), the distributional assumption on the stochastic terms of the marital payoff functions is such that the choice model (2.3.1) is logit and benefits from the property of the independence of irrelevant alternatives. Under this assumption, treating the composition of the married population as exogenously given does not affect the empirical findings.

that is new to this literature. The framework introduced by [Dupuy and Galichon \(2014\)](#) not only allows us to disentangle preferences and demographics effectively, but also to work in a multidimensional and continuous setting. This flexible specification presents an advantage with respect to previous works in that it allows us to analyze different dimensions of sorting at once, in order to understand to what extent marital preferences matter to explain inequality rise, and which dimensions actually have contributed the most to such increase. On the other hand, we limit ourselves to document the changes in sorting patterns and household dynamics without attempting to explain the driving forces behind such transformations. Our work is thus complementary to the richer theoretical frameworks proposed by [Fernández et al. \(2005\)](#) and [Greenwood et al. \(2016\)](#).

Throughout our paper, we provide a detailed picture of the evolution of marital preferences in the United States over the period 1964-2017. In line with the majority of previous works, we find that, even after including several other personal traits, positive assortative mating on education has become stronger and stronger over time. At the same time, positive assortative mating on age has decreased and household specialization seems to be weaker. We also find that, overall, the relevance of socio-economic observable traits has decreased on the marriage market. Finally, preference for racial homogamy seems to have increased since the 1970s: this results seems to be driven by Whites and Hispanics - the fastest growing ethnic minority - growing less and less attracted to each other on the marriage market.

In the second part, we run counterfactual experiments to assess the impact of the shifts in marital preferences on between-household income inequality. We find that, had preferences not changed since 1971, the Gini coefficient would have been 6% lower: this implies that about 25% of the rise of income inequality over the period 1971-2017 is due to changes in sorting patterns. Our results complement those of [Greenwood et al. \(2014\)](#), [Eika et al. \(forthcoming\)](#) and [Greenwood et al. \(2016\)](#) on educational assortativeness. While we find that shifts in marital preferences do matter, we show that they only account for a significant but limited share of the inequality rise. Finally, when decomposing the contribution of marital preferences to the increase of the Gini coefficient, we find that changes in interactions among labor market traits can explain a large share of it. Since the 1980s, couples exhibit a weak but significant complementarity in wage rates and hours worked: this, jointly with important changes in the wage distribution, has crucially contributed to the rise of income inequality. The increased complementarity of spouses' education is also a factor, although the decreased relevance of socio-economic observables in matching and the decreased complementarity of spouses' age are sufficient to offset its effect.

2.A. Technical appendix

2.A.1. Neutrality of optimal matching. According to DG, the equilibrium matching is described by the function 2.3.2. Take the log of $\pi(x, y)$ so that

$$\log \pi(x, y) = x' \frac{A}{\sigma} y - \frac{a(x)}{\sigma} - \frac{b(y)}{\sigma}. \quad (2.A.1)$$

The first component is $x'By$: without the identification assumption with multiple markets described in Section (2.3.4), we are still able to identify $B = A/\sigma$ unequivocally. In fact, under any assumption to disentangle A from σ , a sample (X, Y) yields a unique estimate \hat{B} .

As concerns the second and third components $a(x)/\sigma$ and $b(x)/\sigma$, define $\tilde{a}(x) = \exp(a(x)/\sigma)$ and $\tilde{b}(y) = \exp(b(y)/\sigma)$. We can rewrite (2.3.2) as

$$\pi(x, y) = K(x, y; B) \tilde{a}(x) \tilde{b}(y) \quad (2.A.2)$$

and plug it into the accounting constraints:

$$\begin{aligned} f(x) &= \tilde{a}(x) \int_{\mathcal{Y}} K(x, y; B) \tilde{b}(y) dy \\ f(y) &= \tilde{b}(y) \int_{\mathcal{X}} K(x, y; B) \tilde{a}(x) dx. \end{aligned}$$

DG suggest solving this system by means of an IPFP algorithm. Notice that, we can conclude that, for a given set of parameters B , there is a unique solution given by vectors \tilde{a}^* and \tilde{b}^* , and thus a unique solution π^* .

2.A.2. Improvements to the estimation procedure. Depending on the year, our samples may contain many individuals. However, for computational reasons, estimation can only be performed on a subset of the population. Doing so, we do not make full use of the data to compute the empirical variance-covariance matrix. If the subsample's size is too small, this may even introduce some bias in the estimates. Since the estimation strategy relies on matching the theoretical co-moments to the empirical counterparts, we pick a random subsample whose co-moments of interest are close to those of the full sample. Hence, we use the following procedure to select the subsamples for a dataset with N of observations (one observation corresponds to one couple).

- Step 0. Compute the empirical variance-covariance $\hat{V} \equiv E[XY]$ with the full sample
- Step 1. Draw a subsample of size $n < N$ and compute the empirical variance covariance matrix \hat{V}_n
- Step 2. Check if $|\hat{V} - \hat{V}_n| < \epsilon \times \hat{V}$ for a given level of precision ϵ .

- Step 3. If Step 2 is satisfied, use V_n and the corresponding subsample to estimate the affinity matrix. Otherwise repeat Step 1-3.

2.B. Samples

Table 2.1. Robustness Checks

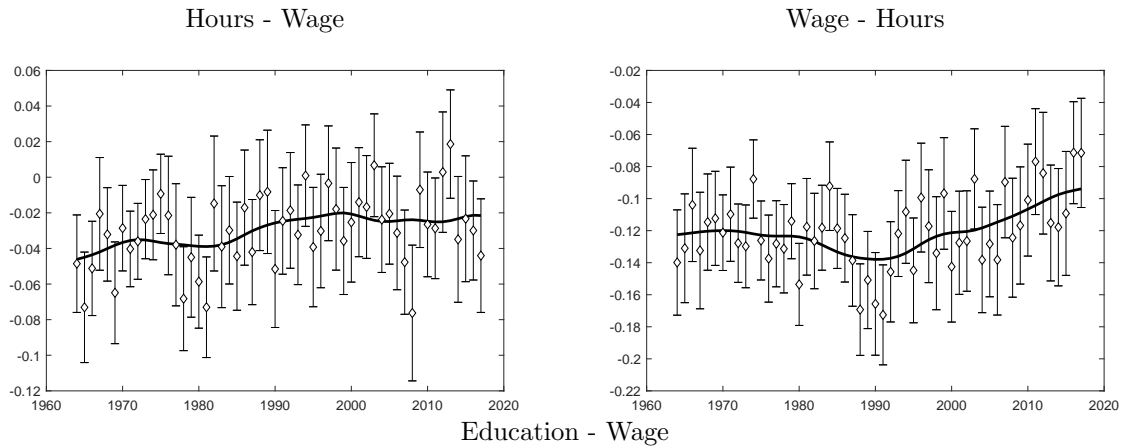
Check #	Period	Age	Education	Wage	Hours	Race	Comments
1. Baselines							
Baseline A	1964-	[23-35] ²⁷	Continuous	Trimmed	Yes	White (incl. Hisp.) and Black	Fig. 2.5,2.6,2.7,2.8, 2.9
Baseline B	1971-	[23-35]	Continuous	Trimmed	Yes	White, Black, Hisp. and Others	Fig. 2.9
Baseline C	1988-	[23-35]	Continuous	Trimmed	Yes	White, Black, Hisp., Asians	Fig. 2.9
2. Robustness checks							
1	1964-	Around median age	Continuous	Trimmed	Yes	White (incl. Hisp.) and Black	Fig. 2.16
2	1964-	[23-35]	Continuous	Trimmed; Annual earnings	Yes	White (incl. Hisp.) and Black	Fig. 2.17
3	1964-	[23-35]	Continuous	Trimmed; Positive	Yes	White (incl. Hisp.) and Black	Fig. 2.18
4	1968-	[23-35]	Continuous	Trimmed;	Yes	White (incl. Hisp.) and Black	Childless couples. Fig. 2.19
5	1964-	[23-35]	Continuous	Trimmed; Potential wage	No	White (incl. Hisp.) and Black	Fig. 2.20
6	1971-	[23-35]	Continuous	Trimmed	Yes	White, Black, Hisp. and Others	One dummy var. per racial category; Fig. 2.21
7	1964-	[23-35]	Excluded	Trimmed	Yes	White (incl. Hisp.) and Black	Fig. 2.22
8	1995-	[23-35]	Continuous	Trimmed	Yes	White (incl. Hisp.) and Black	Cohabiting couples. Fig. 2.23

Note: Bold indicates main changes compared to the baseline.

Notes: the table describes the criteria used to select 11 different samples starting from the main CPS database. These samples are used to run robustness checks throughout the paper.

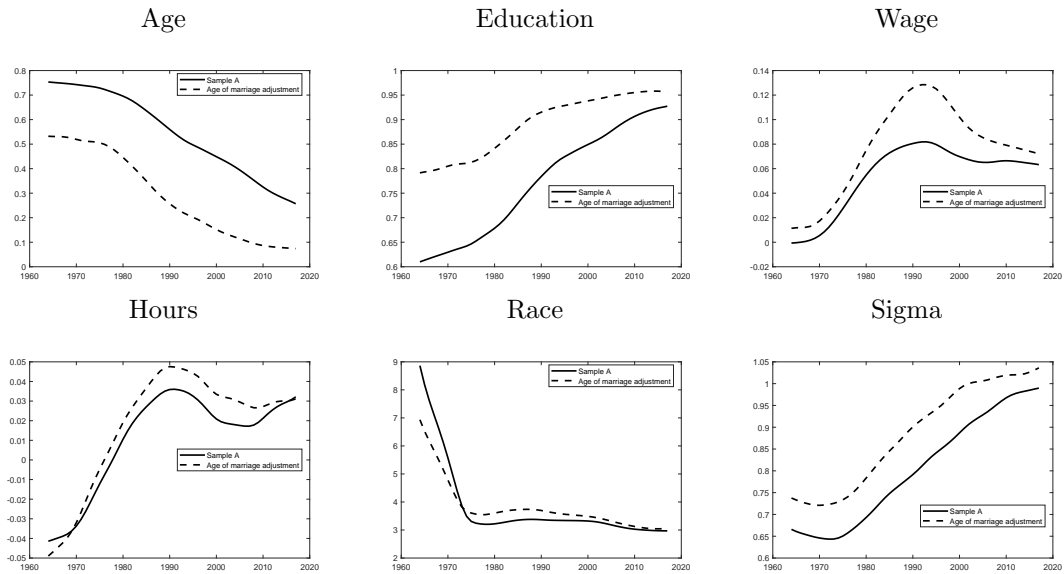
2.C. Additional Figures

Figure 2.15. Relevant cross-interactions



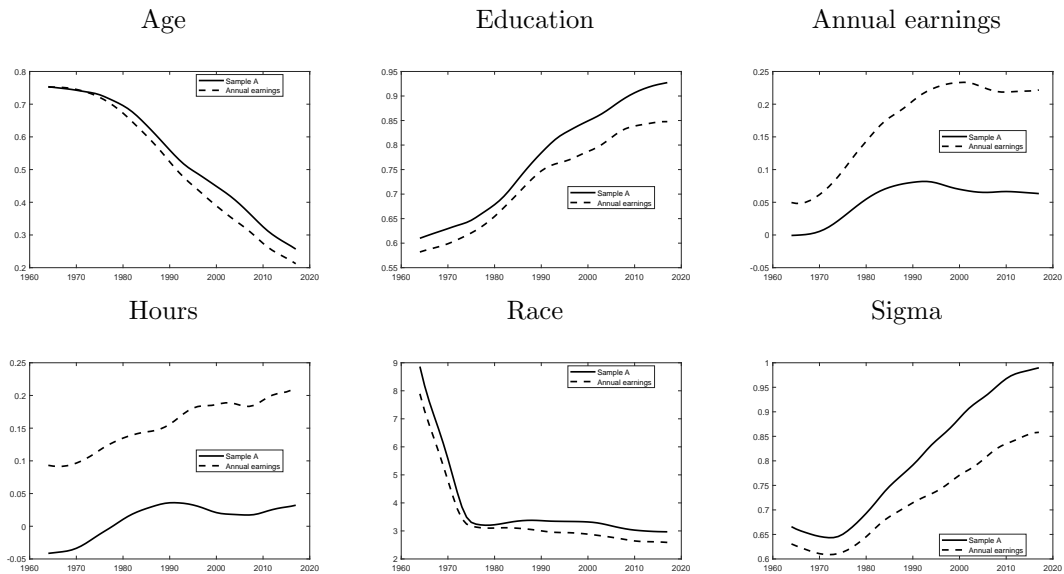
Notes: sample used: baseline A. The figures display the estimated trends of the off-diagonal elements of the marital preference parameter matrix A that have some relevance in our decomposition exercise in Section 2.6.4. In the labels, the first trait is the husband's and the second is the wife's, e.g. Wage-Educ refers to the interaction between the husband's wage rate and the wife's education.

Figure 2.16. Sample selection by median age



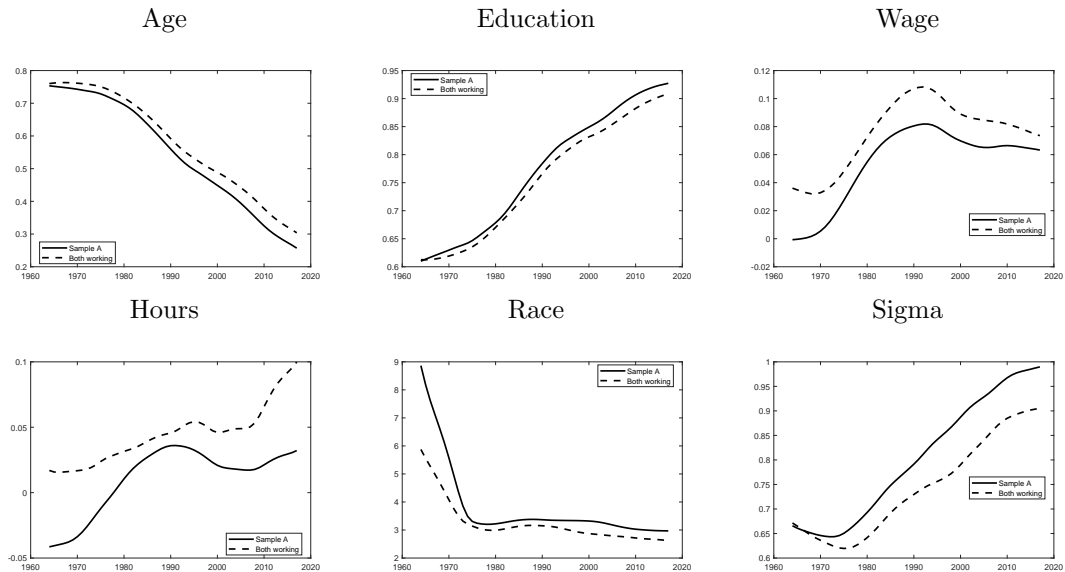
Notes: sample used: check 1 (see Appendix 2.B). The figures compare our baseline results on estimated trends of the diagonal elements of the marital preference parameter matrix A with those obtained using a subsample of couples where at least one partner is aged between the contemporaneous female median age at first marriage (minus 2) and the contemporaneous male median age at first marriage (plus 2).

Figure 2.17. Sorting on annual earnings



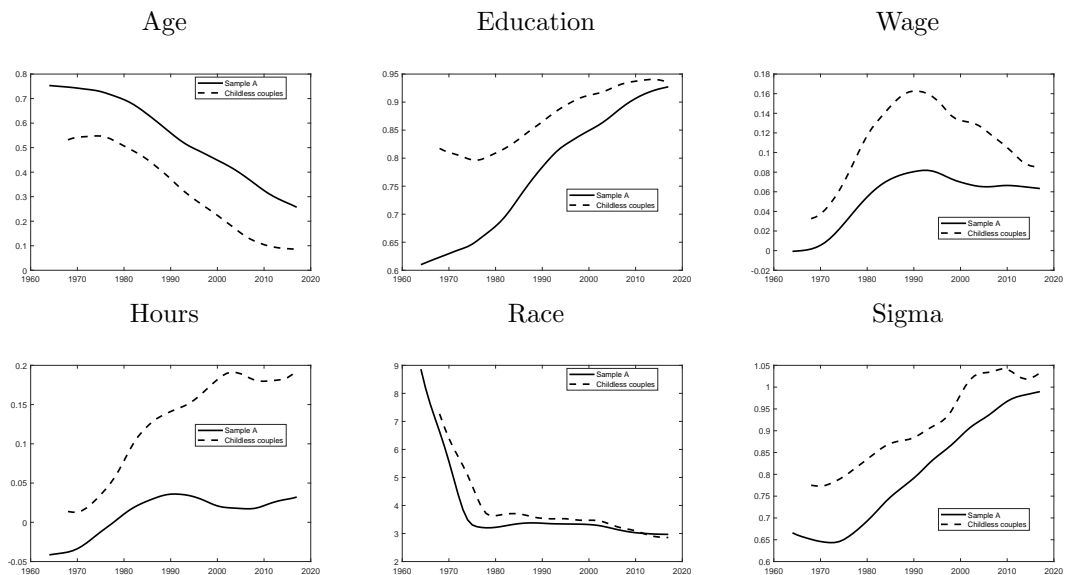
Notes: sample used: check 2 (see Appendix 2.B). The figures show estimated trends of the diagonal elements of the marital preference parameter matrix A when annual earnings are used instead of wages.

Figure 2.18. All couples and working couples (where both partners work)



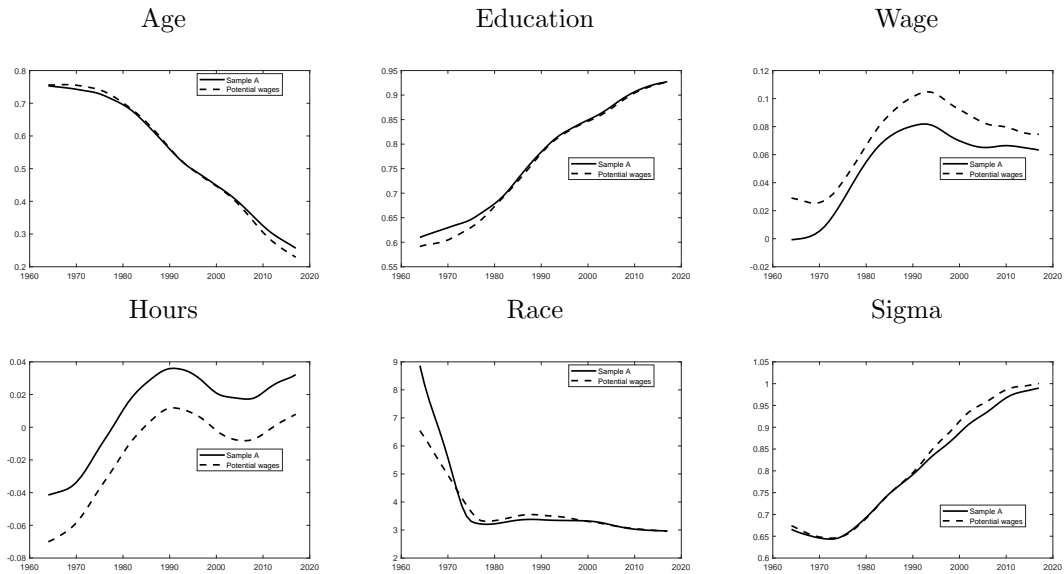
Notes: sample used: check 3 (see Appendix 2.B). The figures compare our baseline results on estimated trends of the diagonal elements of the marital preference parameter matrix A with those obtained using a subsample of couples where both spouses work.

Figure 2.19. All couples and childless couples



Notes: sample used: check 4 (see Appendix 2.B). The figures compare our baseline results on estimated trends of the diagonal elements of the marital preference parameter matrix A with those obtained using a subsample of childless couples.

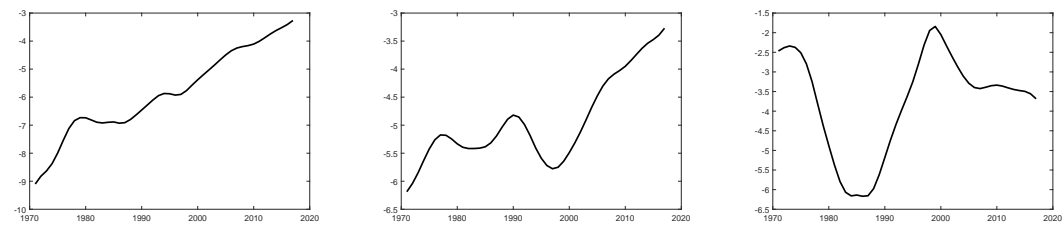
Figure 2.20. Sorting on potential income



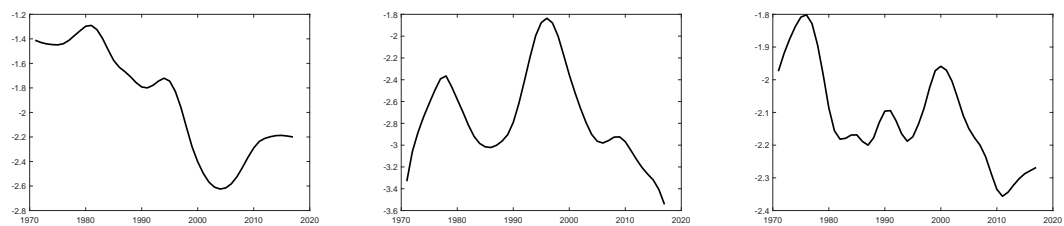
Notes: sample used: check 5 (see Appendix 2.B). The figures compare our baseline results on estimated trends of the diagonal elements of the marital preference parameter matrix A with those obtained using a measure of potential wages.

Figure 2.21. Decomposed taste for interracial marriage

Black Husband, Hispanic Wife Black Husband, White Wife Hispanic Husband, Black Wife

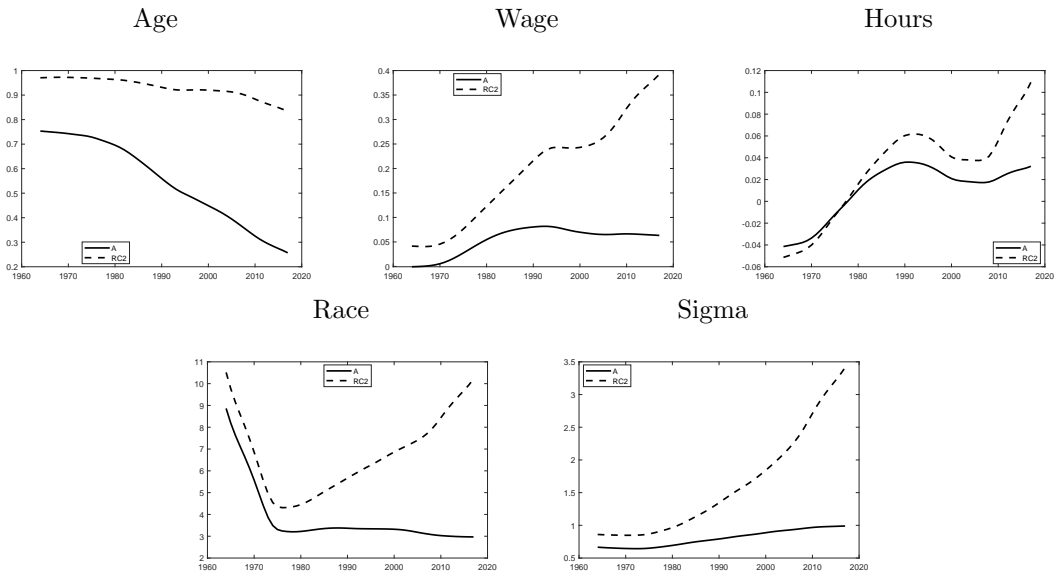


Hispanic Husband, White Wife White Husband, Black Wife White Husband, Hispanic Wife



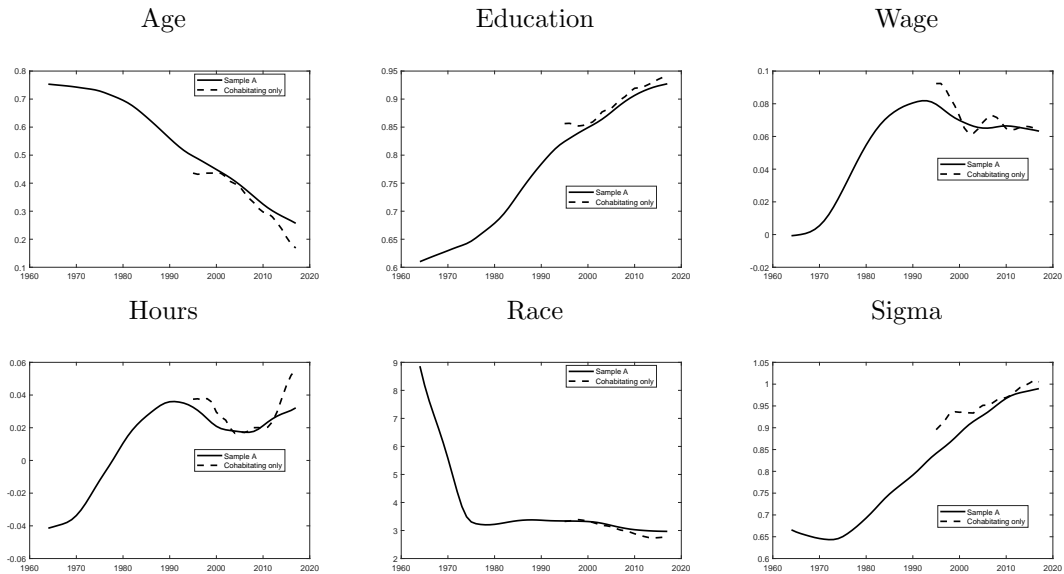
Notes: sample used: check 6 (see Appendix 2.B). The figures describe trends for preferences for interracial marriage, decomposed by racial/ethnic category. We omitted findings for the category “Others”, which is by far the smallest in numbers (see figure 2.3).

Figure 2.22. Education omitted



Notes: sample used: check 7 (see Appendix 2.B). The figures compare our baseline results on estimated trends of the diagonal elements of the marital preference parameter matrix A with those obtained when purposefully omitting education.

Figure 2.23. Cohabiting and married couples



Notes: sample used: check 8 (see Appendix 2.B). The figures compare our baseline results on estimated trends of the diagonal elements of the marital preference parameter matrix A with those obtained using a subsample of cohabiting couples. Data on cohabiting couples are only available since 1995 and the size of the sample is considerably smaller.

Bibliography

- Agresti, A. (2013). *Categorical Data Analysis, 3rd Edition*, John Wiley and Sons.
- Atkinson, M. P. and Glass, B. L. (1985). Marital age heterogamy and homogamy, 1900 to 1980, *Journal of Marriage and the Family* **47**(3): 685–691.
- Becker, G. S. (1973). A theory of marriage: Part i, *Journal of Political Economy* **81**(4): 813–846.
- Becker, G. S. (1974). A theory of marriage: Part ii, *Journal of Political Economy* **82**(2): S11–S26.
- Becker, G. S. (1991). *A Treatise On The Family*, Harvard University Press.
- Blackwell, D. L. (1998). Marital homogamy in the united states: the influence of individual and paternal education, *Social Science Research* **27**(2): 159–188.
- Browning, M., Chiappori, P.-A. and Weiss, Y. (2014). *Economics of the Family*, Cambridge University Press.
- Burtless, G. (1999). Effects of growing wage disparities and changing family composition on the us income distribution, *European Economic Review* **43**(4): 853–865.
- Cancian, M. and Reed, D. (1998). Assessing the effects of wives' earnings on family income inequality, *Review of Economics and Statistics* **80**(1): 73–79.
- Chiappori, P.-A., Salanié, B. and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium, *American Economic Review* **107**(8): 2109–67.
- Choo, E. and Siow, A. (2006). Who marries whom and why, *Journal of Political Economy* **114**(1): 175–201.
- Ciscato, E., Galichon, A. and Goussé, M. (forthcoming). Like attract like? a structural comparison of homogamy across same-sex and different-sex households, *Journal of Political Economy* .
- Dupuy, A. and Galichon, A. (2014). Personality traits and the marriage market, *Journal of Political Economy* **122**(6): 1271–1319.
- Dupuy, A. and Weber, S. (2018). Marital patterns and income inequality, *Technical report*.
- Eika, L., Mogstad, M. and Zafar, B. (forthcoming). Educational assortative mating and household income inequality, *Journal of Political Economy* .
- Fernández, R., Guner, N. and Knowles, J. (2005). Love and money: A theoretical and empirical analysis of household sorting and inequality, *Quarterly Journal of Economics* **120**(1): 273–344.
- Fernández, R. and Rogerson, R. (2001). Sorting and long-run inequality, *Quarterly Journal of Economics* **116**(4): 1305–1341.
- Fryer, R. G. (2007). Guess who's been coming to dinner? trends in interracial marriage over the 20th century, *The Journal of Economic Perspectives* **21**(2): 71–90.
- Fu, X. and Heaton, T. B. (2008). Racial and educational homogamy: 1980 to 2000, *Sociological Perspectives* **51**(4): 735–758.
- Galichon, A. and Salanié, B. (2015). Cupids invisible hand: Social surplus and identification in matching models, *Technical report*, SSRN Working Paper Series.

- Greenwood, J., Guner, N., Kocharkov, G. and Santos, C. (2014). Marry your like: Assortative mating and income inequality, *American Economic Review (Papers and Proceedings)* **104**(5): 348–353.
- Greenwood, J., Guner, N., Kocharkov, G. and Santos, C. (2016). Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation, *American Economic Journal: Macroeconomics* **8**(1): 1–41.
- Heckman, J. J. (1979). Sample selection bias as a specification error, *Econometrica* **47**(1): 153–161.
- Johnson, R. A. (1980). Religious assortative mating in the united states, *Academic Press*.
- Kalmijn, M. (1991a). Shifting boundaries: Trends in religious and educational homogamy, *American Sociological Review* **56**(6): 786–800.
- Kalmijn, M. (1991b). Status homogamy in the united states, *American journal of Sociology* **97**(2): 496–523.
- Kremer, M. (1997). How much does sorting increase inequality?*, *The Quarterly journal of economics* **112**(1): 115–139.
- Liu, H. and Lu, J. (2006). Measuring the degree of assortative mating, *Economics Letters* **92**(3): 317–322.
- Mare, R. D. (1991). Five decades of educational assortative mating, *American Sociological Review* **56**(1): 15–32.
- Qian, Z. and Preston, S. H. (1993). Changes in american marriage, 1972 to 1987: Availability and forces of attraction by age and education, *American Sociological Review* **58**(4): 482–495.
- Schmitt, J. (2003). Creating a consistent hourly wage series from the current population survey's outgoing rotation group, 1979-2002, *Technical report*, CEPR Working Paper.
- Schoen, R. and Wooldredge, J. (1989). Marriage choices in north carolina and virginia, 1969-71 and 1979-81, *Journal of Marriage and the Family* **51**(2): 465–481.
- Schwartz, C. R. (2010). Earnings inequality and the changing association between spouses earnings, *American Journal of Sociology* **115**(5): 1524.
- Schwartz, C. R. and Graf, N. L. (2009). Assortative matching among same-sex and different-sex couples in the united states, 1990–2000, *Demographic Research* **21**(28): 843.
- Schwartz, C. R. and Mare, R. D. (2005). Trends in educational assortative marriage from 1940 to 2003, *Demography* **42**(4): 621–646.
- Shapley, L. S. and Shubik, M. (1971). The assignment game i: The core, *International Journal of Game Theory* **1**(1): 111–130.
- Siegel, J. S. and Swanson, D. A. (2004). *The Methods and Materials of Demography*, Elsevier Academic Press, California, USA.
- Smits, J., Ultee, W. and Lammers, J. (1998). Educational homogamy in 65 countries: An explanation of differences in openness using country-level explanatory variables, *American Sociological Review* **63**(2): 264–285.

- Stevens, G. and Schoen, R. (1988). Linguistic intermarriage in the united states, *Journal of Marriage and the Family* **50**(1): 267–279.
- Stevenson, B. and Wolfers, J. (2007). Marriage and divorce: Changes and their driving forces, *The Journal of Economic Perspectives* **21**(2): 27–52.
- Western, B., Bloome, D. and Percheski, C. (2008). Inequality among american families with children, 1975 to 2005, *American Sociological Review* **73**(6): 903–920.

CHAPTER 3

The changing wage distribution and the decline of marriage

Abstract

In the last forty years, the share of married adults has declined in the United States. At the same time, the structure of labor market earnings has greatly changed, both in its cross-sectional distribution and in terms of life-cycle dynamics. In this paper, I estimate a novel equilibrium model of the marriage market characterized by search frictions, endogenous divorce, aging, and wage mobility. This structural approach allows me to provide a quantitative assessment of the impact of changes in the wage distribution on the decline of marriage. The model aims to rationalize both the cross-sectional marriage patterns - *who marries, and with whom* - and the hazard of marriage and divorce for different subgroups of the population. I find that changes in the wage distribution can account for about a third of the decline in the share of married adults between the 1970s and the 2000s, and partly explain why the decline has been stronger among the low educated. Worsened labor market conditions result in widespread welfare losses among men: these are amplified through the marriage market. Women enjoy welfare gains due to rising wages, although these are partly offset by welfare losses on the marriage market. Increased inequality among men and the shrinking gender wage gap have caused the gains from marriage to shift from household specialization to the possibility of joining efforts on the labor market.

3.1. Introduction

In the last few decades, Americans have faced major changes in the structure of their labor market earnings. In particular, wage inequality has increased both between and within educational groups ([Acemoglu and Autor, 2011](#)). Coincidentally, Americans have also experienced a radical transformation of their family life: the share of married adults has steadily declined since the 1970s, and a decomposition of this trend reveals that the decline has been larger for the low educated.

The relationship between these two synchronous trends has drawn a great deal of attention in the economic literature. Many authors have stressed that increased wage dispersion -

and particularly increased college premium - results in amplified income inequality across households due to the way people sort on the marriage market. However, only few papers have discussed how the marriage market adjusts when we consider changes both in the cross-sectional inequality and in the life-cycle dynamics of labor market earnings. In particular, the structure of wage dynamics along the life-cycle may affect the expected duration of marital relationships and singlehood spells. This is a highly pertinent economic issue as being exposed to longer singlehood spells - either before the first marriage or after a divorce - can bear significant welfare implications.

In this paper, I provide a quantitative assessment of the adjustments of marriage and divorce patterns spurred by the changes in the wage distribution that occurred between the 1970s and the 2000s. To the best of my knowledge, this paper is the first to analyze how the marriage market equilibrium responds to changes not only in wage inequality across education and gender groups, but also in the age profile of wages and in the degree of wage mobility.¹ The empirical analysis focuses on changes in the extensive margin of the marriage market, i.e., the decline of married adults. In particular, the paper aims to address the following questions: *to what extent can changes in the wage distribution explain the overall decline of marriage? How have these changes affected different subgroups of the population, i.e., the high vs the low educated or the younger vs the older? And, finally, how have these changes reshaped the distribution of welfare?*

In order to answer these questions, I build an empirically tractable search-and-matching model of the marriage market where the motives for getting married are both economic and noneconomic. The economic gains from marriage stem from both household specialization and economies of scale. However, along the life-cycle, risk-averse agents face wage shocks and random changes in the quality of the relationship: married couples benefit from the intrinsic insurance mechanism provided by the marriage contract, but, in lack of full commitment, uncertainty may as well lead them to break up. After a divorce, agents are free to marry again, although their prospects change with time due to age and mobile wages. On the aggregate level, systematic differences in matching behavior across educational groups arise as the result of the complex interplay between the primitives of the model: these include the structure of earnings, the home technology, the taste for homogamy, and the way single people meet each other. When taken to the data, the model can rationalize both the cross-sectional marital patterns - *who is married, and with whom* - and the odds of getting married or divorced at different stages of the life-cycle for different subgroups of the population.

This structural approach is motivated by the two following considerations: first, a model is needed to single out the role of changes in labor market earnings in explaining the decline

¹Kopczuk, Saez, and Song (2010) suggest that a natural measure of mobility is the rank correlation in earnings from year t to year $t + p$: I focus on hourly wage mobility in the empirical analysis, and employ an analogous measure - wage rank correlation - to characterize mobility.

of marriage, as opposed to other confounding factors. This is particularly important when comparing marital patterns across an extended timespan, as many factors of different kinds - including demographic, technological, legal and cultural factors - change simultaneously. Second, a model is needed to derive the welfare implications of changes in the expected length of singlehood and marriage spells. In this way, it is possible to identify the population subgroups that have experienced the largest welfare losses.

The empirical analysis relies on the following steps. I first extend the empirical strategy of [Goussé, Jacquemet, and Robin \(2017b\)](#) and provide a formal discussion on how to separately identify the key unobserved parameters of the model: the meeting function, the domestic production function and the cost of divorce. I estimate the model at its stationary state using data moments from both the Current Population Survey (CPS) and the Panel Study of Income Dynamics (PSID) for the American population aged between 20 and 60: the estimation is performed on two separate samples for the 1970s and for the 2000s. I then simulate a series of counterfactual equilibria in order to isolate the role played by changes in a primitive parameter - e.g., the wage distribution - in the transition between the 1970s to the 2000s.

I find that changes in the wage distribution of both men and women can jointly induce a 5.3-percentage-point decrease in the share of married adults. Hence, they account for about a third of the overall decline of marriage between the 1970s and the 2000s. In particular, I show that the shrinking gender wage gap has eroded the economic gains from household specialization for a large part of the population. Most women experience improved labor market conditions over the life-cycle: they can afford being more selective when choosing a partner and value noneconomic gains more than in the past. This leads to the overall decline of marriage.

Changes in the wage distribution can also partly explain why the share of married adults among the low educated has decreased more than among the high educated. The decrease induced by the changing wage distribution is 6pp for men without a college degree and only 2.8pp among male college graduates. Men without a college degree and at the bottom of the wage distribution suffer from the strongest wage cuts: these individuals are thus less suitable for a role of primary earner in a “traditional” household that relies on specialization. On the other hand, high-wage men retain their comparative advantage within the couple over low-wage women and can still enjoy the gains from specialization. Alternatively, they can form couples of two earners with high-wage women, as the spouses’ pooled labor income is sufficiently large to take full advantage of the economies of scale and make the match convenient.

The rich characterization of the wage distribution allows me to delve further into these findings. While I document an increase in wage mobility among men and a decrease among women, I show that - holding wage levels constant - changes in mobility only have

a small impact on the marriage market outcome. I argue that this is due to the estimated high cost of divorce, which indicates that marriage contracts imply a strong commitment, both in the 1970s and in the 2000s. In other words, spouses are likely to insure each other when a labor market shock hits. On the other hand, this implies that individuals are very selective in the choice of the partner, and that adjustments in the equilibrium shares of married adults occur on the entry side: in the 2000s, singles are more selective and search longer before getting married with respect to the 1970s.

Finally, the structural model allows me to quantify the implications of changes in wages on the distribution of welfare and on the gains from marriage. I build intertemporal measures of welfare that account for both the expected stream of labor market earnings during the life-cycle and the expected timing, duration and characteristics of marriage market matches. I show that the improvement of women's labor market conditions with respect to the 1970s leads to welfare gains among women. However, these welfare gains are offset by the lower expected marriage surplus due to worsening marriage prospects over the life-cycle. In addition, men that experience the largest welfare losses due to worsening expectations about their working careers are also those who suffer from the largest losses in terms of expected gains from marriage. I conclude by recovering the distribution of welfare for the 2000s: I show that the market has shifted from an equilibrium where the low and the high educated enjoy, on average, the same level of gains from marriage to an equilibrium where individuals with a college degree enjoy more gains. Hence, individuals at the bottom of the wage and schooling distribution have not only lost ground in terms of human wealth with respect to those at the top², but are also relatively less successful in taking advantage of the additional welfare surplus generated by the marriage market. These considerations leave room for redistributive policies that take into account the monetary value of the gains from marriage.³

The paper provides two main contributions. The first is methodological, and consists in a new type of search-and-matching model that is suitable for the empirical analysis of marriage and divorce decisions along the life-cycle. The model is one of the very first of this kind and includes a number of attractive features that extend previous setups, including wage uncertainty, risk aversion, discrete family labor supply choices and marriage across cohorts: the technical contributions are discussed in detail in the following paragraphs. The framework is potentially suitable for several alternative applications, some of which are discussed in the conclusion. The second contribution is empirical, and consists in a detailed decomposition of the transition from the marriage market outcome observed in the 1970s to the one observed in the 2000s, with the purpose of shedding light on the

²Here and throughout the text, human wealth refers to the discounted sum of life-cycle earnings.

³In the current paper, I do not provide a monetary equivalent of the gains from marriage identified through the matching behavior of agents. Throughout the paper, I briefly discuss how to extend the tools proposed by [Chiappori and Meghir \(2014\)](#), i.e., their Money Metric Welfare Index, to the present search-and-matching framework. This is left to future work.

main driving forces behind the changes. With respect to the previous literature, the paper is the first to deal with the implications of changes in the age profile of wages, in wage mobility, and in the way people meet each other on the marriage market.

The theoretical framework employed in this paper builds on the seminal work of [Shimer and Smith \(2000\)](#), who extend the classic assignment problem of [Shapley and Shubik \(1971\)](#) and [Becker \(1973\)](#) to its two-sided search-and-matching version under Transferable Utility. [Chade and Ventura \(2002\)](#) extend it to allow for random shocks to the agent's types. [Wong \(2003\)](#) estimates the model of Shimer and Smith in order to study marital sorting with respect to wages and education. [Goussé, Jacquemet, and Robin \(2017a\)](#), henceforth, GJR) and [Goussé, Jacquemet, and Robin \(2017b\)](#) merge a non-cooperative household model of consumption, housework and labor supply into a similar search framework: they endogenize divorce by introducing random shocks to the quality of the match, and study the relationship between the marriage market, within-couple bargaining power, and labor supply. [Wiebe \(2018\)](#) is the first to introduce aging in a simplified version of their model without wage heterogeneity. The present paper extends the theoretical framework by GJR in several directions, and in particular by introducing wage uncertainty, risk aversion, aging and discrete labor supply at once. Other working papers moving in the same direction are the works of [Flabbi and Flinn \(2015\)](#), who model both the labor and the marriage market equilibrium, and [Shephard \(2018\)](#). The latter also introduces the life-cycle dimension in the search framework, but complements it with savings and human capital accumulation; the paper focuses on the analysis of age asymmetries in marriage behavior: some key differences between this model and the one in the present work, particularly on the commitment device available to households, are discussed throughout the paper.

From a broader perspective, this work is methodologically related to a larger literature on marriage and matching. [Choo and Siow \(2006\)](#) and [Galichon and Salanié \(2015\)](#) discuss the identification of match surplus when the econometrician observes part of the ex-ante heterogeneity but not the match transfers. Applications to the marriage market are becoming more and more common (e.g. [Dupuy and Galichon, 2014](#)). Extensions to the dynamic case - where agents are free to choose the age of marriage - include [Choo and Siow \(2007\)](#); [Choo \(2015\)](#); [Bruze, Svarer, and Weiss \(2015\)](#): the latter endogenize divorce in a similar fashion to GJR, but introduce a duration-dependent cost function. This group of papers bears interesting similarities with the search-and-matching literature outlined in the previous paragraph, particularly in the identification strategy: these similarities will be discussed throughout the paper.

Another strand of literature incorporates elements of search and/or competitive matching in order to discuss the macroeconomic implications of marital sorting. [Aiyagari, Greenwood, and Guner \(2000\)](#) and [Fernández and Rogerson \(2001\)](#) study the relationship between

Positive Assortative Mating (PAM) and intergenerational mobility; [Fernández, Guner, and Knowles \(2005\)](#) focus on the relationship between human capital investment, PAM and household income inequality. [Regalia and Rios-Rull \(2001\)](#) and [Greenwood, Guner, and Knowles \(2003\)](#) set up intergenerational models of household formation and dissolution in order to study how marital patterns, fertility and inequality are jointly determined at equilibrium. [Greenwood, Guner, Kocharkov, and Santos \(2016\)](#) focus on the determinants of rising cross-sectional income inequality: they estimate a dynamic model of educational choice, marriage, divorce, and labor supply, and show that marriage market adjustments as a response to changes in wages can partly account for the increase in the Gini coefficient. The propagation mechanism they describe explains why the share of married adults falls, and is similar to the one suggested in this paper: in particular, both papers stress the importance of increasing wage inequality and of the shrinking gender wage gap. However, in the present paper, I move the focus away from cross-sectional household income inequality and build measures of intertemporal welfare that take into account one's expectations about both his/her future working career and his/her future gains from marriage.

In spite of a much simplified labor supply setting, this paper is also related to a large strand of literature that studies labor supply, savings, fertility and/or childcare decisions along the life-cycle. Building on the seminal work of [Eckstein and Wolpin \(1989\)](#), many works in this literature include marital status as one of the household's state variable, and transitions are modeled as exogenous shocks (e.g. [Eckstein and Lifshitz, 2011](#); [Blundell, Dias, Meghir, and Shaw, 2016](#); [Adda, Dustmann, and Stevens, 2017](#)). Some works focus on marital dissolution and take the initial family composition as given: for instance, [Gemici \(2011\)](#) and [Voena \(2015\)](#) study the determinants of divorce - focusing on geographical mobility and divorce laws respectively - taking the initial household composition as given.

Building on this life-cycle labor supply literature, a growing number of papers considers endogenous marriage and divorce decisions: a precursor is [Van der Klaauw \(1996\)](#), while more recent examples are [Sheran \(2007\)](#), [Keane and Wolpin \(2007\)](#), [Keane and Wolpin \(2010\)](#), [Bronson \(2014\)](#), [Mazzocco, Ruiz, and Yamaguchi \(2017\)](#) and [Fernández and Wong \(2017\)](#). Each of these papers focuses on a specific empirical issue, but all of them feature a rich characterization of household behavior (particularly, of savings and human capital accumulation) and a focus on the welfare implications of changes in earnings, policy parameters and marriage market opportunities. In this paper, I depart from this literature by treating the supply of available partners as an endogenous equilibrium object rather than imposing an exogenous distribution from where to draw candidate partners. Two notable exceptions are the works of [Reynoso \(2017\)](#) and [Beauchamp, Sanzenbacher, Seitz, and Skira \(2018\)](#): the first studies the impact of divorce laws on marriage, labor supply and divorce; the marriage market is thought as a static matching game played at the beginning of adulthood, and remarriage is not allowed. The second paper estimates an

equilibrium model of the marriage market with divorce and remarriage in order to study the determinants of single parenthood: their definition of equilibrium differs from the one used in this paper as agents are only allowed to marry within their own cohort.

The paper is structured as follows: I first present the model in section 3.2; I provide a formal discussion on how to identify its key unobserved parameters in section 3.3; I introduce the CPS and PSID samples in section 3.4; I provide details about the empirical specification and the estimation procedure in section 3.5; I present the results of the estimation, of the counterfactual analysis and of the welfare analysis in section 3.6; I conclude in section 3.7.

3.2. Model

The theoretical framework extends the original two-sided search-and-matching model by [Shimer and Smith \(2000\)](#) in the vein of GJR. Single agents search for possible partners on the marriage market. Married households face two layers of uncertainty: first, about the future quality of the current match; second, about the future wage rates of the spouses. When uncertainty is resolved, the couple needs to make a decision about whether to continue the match or not: in this way, divorce decisions are endogenous.

In contrast with the previous literature, I introduce some key elements that extend the empirical analysis to the life-cycle level. First, I introduce aging: agents get older as time goes by, and age influences their odds of marriage, divorce and remarriage. Second, I assume agents are risk-averse and experience wage shocks along the life-cycle: wage uncertainty partly explains marital instability, as economic gains from marriage may disappear following a wage shock.

Before turning to the setup, let me anticipate some key implications of these assumptions. On the one hand, the definition of a *deterministic* steady-state equilibrium⁴ requires some strong assumptions on market entry and exit. In the stationary environment, new cohorts do not differ from the previous ones in terms of size and composition, and, on the aggregate, display the same matching behavior along the life-cycle: both business cycles and secular drifts are not captured by the model, and only comparative statics can help make sense of differences across time and space. On the other hand, the introduction of aging does not require the wage process to be stationary, nor new marriages to be outbalanced by an equal amount of divorces for a given group of agents. Different groups *within* the same cohort may take strongly diverging paths in terms of both earnings and family achievements, exactly as we see in the data.

⁴As opposed to a *stochastic* steady-state, where equilibrium quantities are functions of an aggregate state of the world. In this search-and-matching model, there are several assumptions that could be relaxed in order to introduce aggregate uncertainty. Some possible extensions will be discussed in this section and in the conclusion.

The section is organized as follows. First, in subsections 3.2.1, 3.2.3 and 3.2.2, I describe the general environment and the *ex-ante* heterogeneity characterizing agents. Then, in subsections 3.2.4, 3.2.5 and 3.2.6, I describe the household problem for couples and singles. Finally, in subsections 3.2.7, 3.2.8, 3.2.9 and 3.2.11, I describe the search environment and marriage and divorce decisions, and, in section 3.2.12, I conclude by providing a definition of the steady-state equilibrium.

3.2.1. Heterogeneous agents and aging. Men are *ex-ante* heterogeneous and are associated with a publicly observable *type* i , a vector comprising the following elements:

- a time-invariant human capital type h_i ;
- age a_i , which is deterministically updated over time;
- current wage w_i , which changes over time according to an $AR(1)$ random process described in the section 3.2.3.

Similarly, the type of a woman is given by $j = (h_j, a_j, w_j)$. The men's (women's) set of types is named \mathcal{I} (\mathcal{J}) and is discrete. Note that the time subscript t is unnecessary, since I will focus on the steady-state equilibrium. Agents care about time because of the aging and stochastic wage process, but live in a stationary environment.

Aging individuals discount future by means of a factor β , and face an exogenous probability of exiting the market. A man i exits the marriage market with probability $1 - \psi_m(i)$ at the end of the period. If he survives, he grows one-year older, so that i' is such that $a_{i'} = a_i + 1$. A similar process governs women's aging, with a different vector of survival probabilities ψ_f . In addition, assume agents eventually leave with probability one at \bar{a} , i.e., $\psi_m(\bar{a}) = \psi_f(\bar{a}) = 0$. Market exit is primarily intended as death, although it could also encompass other situations where the agent is unable to live in a two-adult household nor to look for a partner: these may include active duty in the army, incarceration, long-term stay in health-care institutions, and so on.

3.2.2. Marital status and timing. Time is discrete, and a period is defined as the timespan between t and $t + 1$. In t , an agent is associated with his (her) current type i and (j) and is either married or single. Married couples are characterized by the joint type (i, j) . At the very beginning of the period, in t_+ , uncertainty is resolved. First, agents learn whether they will live on to the next period, with exit probabilities described by ψ_g , $g \in \{m, f\}$. If they do, they grow one-year older and draw new wages according to the $AR(1)$ wage process described in the next section.

Consider the time-line of a given period for a man i and a woman j that are married at the beginning of the period (see figure 3.1). On top of drawing new wages, the couple also observes the realization of a vector $\boldsymbol{\eta}$ of temporary shocks that help characterize the

quality of the current match. The distribution and exact role of $\boldsymbol{\eta}$ will be described in the next section. However, assume from now that the vector $\boldsymbol{\eta}$ is *i.i.d.* across time and couples. Conditionally on their new types i' and j' and on $\boldsymbol{\eta}$, the spouses decide whether they should stay together until the end of the period. If they divorce, they both stay single until the end of the period. Finally, conditionally on their updated marital status, agents make consumption and labor supply choices.

Consider the time-line of a given period for an individual who is single at the beginning of the period (see figure 3.2). In t_+ , she draws a new wage and learns her new type, and she may also meet a someone of the opposite of sex (in the case of women, they could meet a man of type i). The pair draws a vector of shocks $\boldsymbol{\eta}$ that will influence the subsequent matching decision, namely whether to get married or to stay single until the end of the period. Finally, the woman makes consumption and labor supply choices, either alone or with her new husband.

3.2.3. Wage process. Wages follow an *AR*(1) process, so that, at the beginning of every period, a man draws a new wage $w_{i'}$ conditionally on both his current wage w_i and his deterministic traits h_i and a_i . Because of the randomness of the wage process, the probability of transiting from type i to type i' is denoted $\pi_m(i, i') \equiv Pr(i'|i)$. The notation used for $\pi_m(i, i')$ stresses that $\pi_m(i, i')$ depends on the full vector i , i.e., on both the current wage w_i , age a_i and human capital h_i .⁵ Analogous considerations hold for women, whose probability of transiting from j to j' is given by $\pi_f(j, j')$, where π_f is possibly different from π_m .

The wage process does not need to be stationary due to life-cycle dynamics being taken into account. Importantly, for a given cohort of people, mean wages depend on age, and a cohort's wage dispersion may increase along the life-cycle. While I let time-invariant traits and age directly influence the conditional distribution of wages, I do not attempt to decompose wages in multiple factors, and in particular to distinguish between a permanent and a temporary component as it is common in the literature (see [Meghir and Pistaferri, 2011](#)). Therefore, wage shocks need to be interpreted as permanent: I will often refer to *wage mobility* as the extent to what individuals are likely to move along the wage distribution from one period to the next (see [Kopczuk, Saez, and Song, 2010](#)).

Finally, note that the assumption that the wage process is fully exogenous - and in particular that it is not affected by the agent's marital status - is likely to be highly counterfactual, particularly for women. Joint household labor supply decisions may have an impact on human capital accumulation, especially if household specialization plays

⁵Note that the corresponding transition matrix $(\pi_m(i, i')) \in \mathbb{R}^{|\mathcal{Z}| \times |\mathcal{Z}|}$ is such that $\pi_m(i, i') = 0$ if $h_{i'} \neq h_i$ or if $a_{i'} \neq a_i + 1$.

an important role as a motive to marry. Hence, one's marital status may influence the evolution of his/her wage rate. These crucial limitations are discussed in the conclusion.

3.2.4. Household problem: preferences and domestic production. With timelines 3.1 and 3.2 in mind, it is possible to characterize agents' rational behavior by solving backwards for their optimal choices. In this section, I will introduce agents' preferences, while in sections 3.2.5 and 3.2.6 I will solve the household problem. Only starting from section 3.2.7, I will proceed by describing the optimal matching decisions (i.e., marriage and divorce).

Agents enjoy utility from the consumption of both a private good q and a public good Q . The agents' utility is represented by the following function:

$$u(q, Q) = \frac{1}{2} \log q + \frac{1}{2} \log Q. \quad (3.2.1)$$

The good Q can be thought of as intermediary and is produced domestically using both time and money input. The production function of Q is

$$Q = \begin{cases} (t_m + t_f)^{\gamma_1(i,j)} \exp(\gamma_2(l_f; j) + \gamma_3(i, j) + \eta_{l_f}) & \text{for married households} \\ t_m^0 & \text{for single men} \\ t_f^0 & \text{for single women.} \end{cases} \quad (3.2.2)$$

For married households, t_m and t_f represent the husband's and wife's share of public good expenditure, assumed to be perfect substitutes. The elasticity of Q with respect to the total expenditure is equal to γ_1 , and may depend on the spouses' characteristics: γ_1 plays an important role in the empirical analysis of marriage, in that it determines the size of the economies of scale enjoyed by married couples. As concerns single adults, they can only produce Q via a monetary input: t_g^0 is the share of their budget allocated to home production.⁶

The total amount of time available to an agent is normalized to one, so that $l_f \in \mathcal{L}_f$ represents the wife's share of time spent on the labor market, where \mathcal{L} is assumed to be discrete. Married men always spend the entire time endowment on the labor market: while the theoretical framework does not require the labor supply of men to be fixed, restricting the choice set in this direction is convenient for the empirical analysis and is broadly consistent with the patterns observed in the data.⁷ There exist possible public benefits of having a stay-at-home spouse, which are captured by the productivity shifter γ_2 . In other words, if γ_2 is decreasing in l_f , the couple benefits from a higher public good

⁶The elasticity of Q with respect to t_g^0 is normalized to 1. The role of this normalization to ensure the identification of the model is remarked in section 3.3.

⁷In the CPS, only about 10% of married men between 20 and 60 are out of the labor force in the 1970s, 8% in the 2000s.

production when the spouse reduces her labor market effort, all else constant. In practice, γ_2 is left unrestricted and is estimated for different levels l_f in the empirical analysis.

The term γ_3 is an additional productivity shifter, which depends on the couple's type (i, j) only. It has the role of capturing the relevance of additional interactions between traits, such as a preference for educational homogeneity or age proximity. Finally, the term η_{l_f} is a random shifter taken from the vector $\boldsymbol{\eta} \in \mathbb{R}^{|\mathcal{L}_f|}$: each option l_f is associated with an element of $\boldsymbol{\eta}$, and each η_{l_f} is distributed logistically with location and scale parameters normalized to 0 and 1 respectively. While $\boldsymbol{\eta}$ is *i.i.d.* across time and couples, its elements may be correlated with each other, with $0 < \sigma_\ell \leq 1$ representing their degree of independence (see Train, 2009, Chapter 4, for details, and note that with $\sigma_\ell = 1$ we have a standard logit model).

In the case of married households, an allocation $(q_m, q_f, t_m, t_f, l_f)$ is feasible if the following private budget constraints are respected:

$$q_m = w_i - t_m \quad (3.2.3)$$

$$q_f = l_f w_j - t_f \quad (3.2.4)$$

where the sign of t_m and t_f is unrestricted. Hence, $t_m < 0$ ($t_f < 0$) implies that the wife (husband) is actually transferring money into the husband's (wife's) pocket.

Analogous budget constraints hold for single agents. However, it is assumed hereafter that singles always work full-time. Since the household problem is fully static, the labor supply of singles does not have an impact on their future marriage choices, and is thus not included in the analysis.⁸

3.2.5. Household problem: public good. As a first step to solve the household problem for married agents, assume the optimal household allocation $(q_m^*, q_f^*, t_m^*, t_f^*, l_f^*)$ is efficient. This assumption puts the model in the general collective framework introduced by Chiappori (1988, 1992). Moreover, the preferences implied by the utility function (3.2.1) verify the Transferable Utility (TU) property.

The efficiency assumption and the TU property bear two important implications. First, the demand for public good $Q^*(l_f; i, j)$ conditional on the wife's labor supply l_f does not depend on the point of the Pareto frontier chosen by the household. On the other hand, the couple may disagree on the repartition (t_m, t_f) : this will be dealt with in the next section. Second, the spouses always agree on the level l_f that puts them on the outermost Pareto frontier. This second statement relies on the absence of private incentives (e.g.,

⁸Allowing singles to adjust their labor supply is possible, and would be crucial if human capital accumulation choices were considered. However, it implies a slightly more complicated discrete choice setting. Shephard (2018) shows a convenient way of handling larger discrete choice sets by partly postponing the resolution of uncertainty after the matching phase.

private leisure) for the wife to manipulate her supply l_f : in this particular case, the $|\mathcal{L}|$ Pareto frontiers are parallel, and there is no disagreement about how much the wife should work. The underlying economic intuition is that, while the wife might be “unhappy” about the selected l_f , she can always be compensated through a more favorable division of the public good expenditure.

In order to derive the demand for public good, substitute the budget constraints (3.2.3) and (3.2.4) into the utility function (3.2.1) and define the *conditional indirect utilities* as follows:

$$\begin{aligned}\phi_m(t_m, t_f, l_f; i, j) + \frac{\eta_f}{2} &= \frac{1}{2} \log(w_i - t_m) + \frac{\gamma_1(i, j)}{2} \log(t_m + t_f) + \frac{\gamma_2(l_f; j)}{2} + \frac{\gamma_3(i, j)}{2} + \frac{\eta_f}{2} \\ \phi_f(t_m, t_f, l_f; i, j) + \frac{\eta_f}{2} &= \frac{1}{2} \log(l_f w_j - t_f) + \frac{\gamma_1(i, j)}{2} \log(t_m + t_f) + \frac{\gamma_2(l_f; j)}{2} + \frac{\gamma_3(i, j)}{2} + \frac{\eta_f}{2}.\end{aligned}$$

The demand $Q^*(l_f; i, j)$ is then obtained by maximizing any weighted sum of ϕ_m and ϕ_f with respect to t_m and t_f , holding l_f fixed. The share of household budget used as an input for Q is:

$$\frac{\gamma_1(i, j)}{1 + \gamma_1(i, j)}(w_i + l_f w_j). \quad (3.2.5)$$

It is possible to provide a linear characterization of the Pareto frontier associated with a given level l_f by summing up the exponentials of the indirect utility functions (see [Chiappori and Gugl, 2014](#)). Their sum is equal to a constant, and helps characterize the Pareto set in a way that the Pareto frontier is a straight line with slope -1 :

$$\exp(\phi_m(t_m, t_f, l_f; i, j)) + \exp(\phi_f(t_m, t_f, l_f; i, j)) \equiv \Gamma(l_f; i, j) \quad (3.2.6)$$

where Γ is calculated explicitly in appendix 3.A.1.

When it comes to labor supply decisions, the spouses will choose l_f in order to select the outermost Pareto frontier, once the random shocks $\boldsymbol{\eta}$ are taken into account:

$$l_f^*(i, j) = \arg \max_{l_f \in \mathcal{L}} \Gamma(l_f; i, j). \quad (3.2.7)$$

3.2.6. Household problem: sharing rule. Another key implication of the Transferable Utility property is that it is impossible to recover the *sharing rule* (t_m, t_f) , i.e., the exact point on the Pareto frontier chosen by the spouses. The solution $(t_m^*(l_f; i, j), t_f^*(l_f; i, j))$ can only be pinned down if the distribution of power within the household is known. As suggested by [Becker \(1973\)](#), the sharing rule within the couple responds to shifts in supply and demand of mates of a given type in the marriage market. However, when search frictions are present, it is not possible to uniquely characterize Pareto weights starting from the marriage market outcome ([Shimer and Smith, 2000](#)). The underlying intuition can be explained with this thought example: consider a man who proposes to a woman and offers her a marriage contract that makes her just indifferent between marrying him

and keeping on searching. Due to search frictions, she cannot turn to another similar man and agree on a better deal where she extracts a slightly larger share of surplus. In fact, waiting might be too costly, and, if the surplus is positive, the woman still has an incentive to accept. Yet, the marriage market still plays a crucial role in determining whether there is a set of allocations that make both candidates better off together than singles. If the set is not null, the choice of the allocation also depends on an additional *within-couple bargaining mechanism*. In some markets, the set of feasible allocations can be small, and competition might greatly reduce the role of within-couple bargaining; in others, the reverse can be true.

The discussion above implies that it is necessary to introduce an additional bargaining mechanism in order to recover the sharing rule and close the model. To understand how $(t_m^*(i, j), t_f^*(i, j))$ is selected, let me first introduce the relevant bargaining payoffs. These are given by the present discounted value of all expected flows in the future, once accounted for the possibility of breakups. Define $(W_m(t_m, t_f, l_f; i, j), W_f(t_m, t_f, l_f; i, j))$ as the *present discounted value of marriage* for a man and a woman in a couple (i, j) under sharing rule (t_m, t_f) and with labor supply l_f . In addition, define $(V_m^0(i), V_f^0(j))$ as the respective *present discounted value of singlehood*. The latter constitute the bargaining breakpoint and are shaped by the marriage market: agents take them as given during the bargaining phase.

Call $(t_m^*(l_f; i, j), t_f^*(l_f; i, j))$ the solution to the bargaining process for a couple (i, j) conditional on choosing labor supply l_f . The respective marriage payoffs are

$$V_m(l_f; i, j) + \frac{\eta_{l_f}}{2} \equiv W_m(t_m^*(l_f; i, j), t_f^*(l_f; i, j), l_f; i, j) + \frac{\eta_{l_f}}{2} \quad (3.2.8)$$

$$V_f(l_f; i, j) + \frac{\eta_{l_f}}{2} \equiv W_f(t_m^*(l_f; i, j), t_f^*(l_f; i, j), l_f; i, j) + \frac{\eta_{l_f}}{2}. \quad (3.2.9)$$

I assume that $(t_m^*(l_f; i, j), t_f^*(l_f; i, j))$ are chosen so that the following *surplus splitting rule* holds:

$$V_f(l_f; i, j) - V_f^0(j) = \frac{V_f(l_f; i, j) - V_f^0(j) + V_m(l_f; i, j) - V_m^0(i)}{2} \equiv \frac{S(l_f; i, j)}{2} \quad (3.2.10)$$

where S is the *systematic marriage surplus*, i.e., the total surplus net of the temporary match-quality shock η_{l_f} . In practice, the spouses split the total surplus in equal parts.

The splitting rule assumed in this model is arguably a simple one. Yet, it is able to generate a significant amount of variation across individuals in terms of private consumption. A closed-form equation for individual demand functions q_m^* and q_f^* is derived in appendix 3.A.2. Once the sharing rule has been recovered, it is possible to derive the per-period indirect utilities $v_g(l_f; i, j) \equiv \phi_g(t_m^*(l_f; i, j), t_f^*(l_f; i, j), l_f; i, j)$ for $g \in \{m, f\}$.

In the previous literature, [Shimer and Smith \(2000\)](#) and GJR assume that couples select a sharing rule through Nash bargaining. Collective models are a more general representation of the household problem: in this paper, I restrict *Pareto weights* to 0.5 for all couples. This simple rule has the advantage of delivering a closed-form equation for surplus, and thus to make the computation of the steady-state equilibrium faster.⁹

To conclude this section, the following lemma can now help characterize the wife's labor supply decision.

LEMMA 1. *Assume that ϕ_m and ϕ_f verify the Transferable Utility property and that surplus is split according to rule (3.2.10). Then,*

$$S(l_f; i, j) + \eta_{l_f} > S(l'_f; i, j) + \eta_{l'_f} \Rightarrow \Gamma(l_f; i, j) > \Gamma(l'_f; i, j). \quad (3.2.11)$$

PROOF. It is possible to write $V_m(l_f; i, j) = \phi_m(t_m^*(l_f; i, j), t_f^*(l_f; i, j), l_f; i, j) + C_m(i, j)$ and $V_f(l_f; i, j) = \phi_f(t_m^*(l_f; i, j), t_f^*(l_f; i, j), l_f; i, j) + C_f(i, j)$, where C_m and C_f are, respectively, the continuation values for the husband and the wife. The latter do not depend on current labor supply due to the household problem being static. Consider l_f and l'_f such that $S(l_f; i, j) + \eta_{l_f} > S(l'_f; i, j) + \eta_{l'_f}$, and introduce the notation $\delta f(x) = \delta f(l_f; x) - \delta f(l'_f; x)$, so that $\delta S(i, j) + \delta \eta > 0$. Both the reservation utilities and continuation values are the same regardless of the choice of l_f : hence, the surplus differential is given by $\delta S(i, j) = \delta \phi_m(i, j) + \delta \phi_f(i, j)$. Similarly, the husband's surplus differential is $\delta V_m(i, j) + \delta \eta/2 = \delta \phi_m + \delta \eta/2 = \delta S(i, j)/2 + \delta \eta/2$, where the last equality is due to the splitting rule (3.2.10). Since $\delta S(i, j) + \delta \eta > 0$, both $\delta \phi_m + \delta \eta/2$ and, for analogous reasons, $\delta \phi_f + \delta \eta/2$ are positive. Now, recall that a necessary and sufficient condition for the TU property to hold is the existence of two functions g_m and g_f , continuous and increasing, s.t. the Pareto frontier can be expressed as

$$g_m(\phi_m(t_m, t_f, l_f; i, j) + \frac{\eta_{l_f}}{2}) + g_f(\phi_f(t_m, t_f, l_f; i, j) + \frac{\eta_{l_f}}{2}) = \Gamma(l_f; i, j),$$

a result exposed, e.g., in [Chiappori and Gugl \(2014\)](#) and [Demuyneck and Potoms \(2018\)](#). Due to g_m and g_f being increasing, $g_m(\phi_m(l_f) + \frac{\eta_{l_f}}{2}) > g_m(\phi_m(l'_f) + \frac{\eta_{l'_f}}{2})$ and $g_f(\phi_f(l_f) + \frac{\eta_{l_f}}{2}) > g_f(\phi_f(l'_f) + \frac{\eta_{l'_f}}{2})$. Adding up the two inequalities, one obtains $\Gamma(l_f) > \Gamma(l'_f)$. \square

Lemma 1 establishes that the splitting rule (3.2.10) implies that the household always chooses the level of labor supply associated with the highest total surplus $S(l_f; i, j) + \eta_{l_f}$. The chosen level of labor supply l_f , $(t_m^*(l_f; i, j), t_f^*(l_f; i, j))$ lies on the outermost Pareto

⁹GJR estimate the Nash bargaining parameter, which corresponds to the Pareto weight when utility functions are taken to be quasilinear (which is not the case in the current paper). They find it to be not significantly different than 0.5. Alternatively, the Pareto weight may be allowed to differ across couples, to depend on reservation utilities ($V_m^0(i), V_f^0(j)$) or even on some (time-invariant) *distribution factors* in the spirit of [Browning and Chiappori \(1998\)](#).

frontier if $S(l_f; i, j) + \eta_{l_f} > S(l'_f; i, j) + \eta_{l'_f}$.¹⁰ The proof suggests that, when switching from l'_f to l_f , the additional surplus is always redistributed proportionally, so that the ratio of the shares is always 0.5. Thanks to the additive separability of η_{l_f} and the logit assumption, the probability of selecting a given level l_f can be written as:

$$\begin{aligned} \ell(l_f; i, j) &\equiv \Pr \left\{ S(l_f; i, j) + \eta_{l_f} > S(l'_f; i, j) + \eta_{l'_f} \quad \forall l'_f \in \mathcal{L} \right\} = \\ &= \frac{\exp(S(l_f; i, j)/\sigma_\ell)}{\sum_{l' \in \mathcal{L}} \exp(S(l'; i, j)/\sigma_\ell)}. \end{aligned} \quad (3.2.12)$$

Taking expectations over marriage surplus with respect to the vector $\boldsymbol{\eta}$ yields the following expected surplus (named *inclusive surplus* in the GEV literature; see [Train, 2009](#)):

$$\bar{S}(i, j) \equiv \sigma_\ell \log \sum_k \exp(S(k; i, j)/\sigma_\ell). \quad (3.2.13)$$

3.2.7. Divorce decisions. Up to now, household composition was treated as given. However, when the uncertainty is revolved in t_+ (see timeline 3.1), the couple may decide to break up. Conditionally on the realization of the wage shocks and on temporary shocks $\boldsymbol{\eta}$, spouses are free to compare the optimal household allocation that they can achieve in the coming period to what they can get if each of them were on his/her own. In other words, the new household allocation has to respect the spouses' individual rationality constraints. However, in this model, spouses also face a one-time cost of divorce.

The updated state of the couple in t_+ is given by the spouses' new types (i, j) and the updated vector of match-quality shocks $\boldsymbol{\eta}$. If, once accounted for the new shocks $\boldsymbol{\eta}$, the total surplus is not positive for the outermost Pareto frontier $\Gamma(l_f; i, j)$, then the spouses are better off breaking up. Given the logit assumption on $\boldsymbol{\eta}$, the probability of divorce is given by:

$$\begin{aligned} 1 - \alpha(i, j) &\equiv \Pr \left\{ \max_{l_f \in \mathcal{L}} (S(l_f; i, j) + \eta_{l_f}) < -\kappa \right\} = \\ &= \frac{\exp(-\kappa)}{\exp(-\kappa) + \left(\sum_{l_f} \exp(S(l_f; i, j)/\sigma_\ell) \right)^{\sigma_\ell}} = \frac{1}{1 + \exp(\bar{S}(i, j) + \kappa)}. \end{aligned} \quad (3.2.14)$$

where the last equality follows from the definition of inclusive surplus (3.2.13). The probability $\alpha(i, j)$ corresponds to the odds of continuing the current marriage. Finally, the parameter κ stands for the sunk cost of divorce: note that, *ceteris paribus*, a higher κ leads to a higher α .

¹⁰The caveat is that the sum $\phi_m(l_f) + \phi_f(l_f)$ does *not* necessarily provide an implicit representation of the Pareto frontier. It only does so in the case where ϕ_m (ϕ_f) is quasilinear in the private budget $w_i - t_m$ ($w_j - t_f$): in this case, the functions g_m and g_f required to obtain a linear representation of the Pareto frontiers as in (3.2.6) are actually linear functions.

When divorce cannot be ruled out, only limited commitment devices are feasible: these mechanisms have been widely studied in the economic literature on marriage and risk-sharing. In this model, it is assumed that, when uncertainty is resolved, agents are completely free to bargain over a new household allocation: if bargaining fails because of negative surplus, the couple splits after paying a one-time “fine”. In this model, the only source of commitment is the cost of divorce. However, the fact that married agents are free to rediscuss the sharing rule following any shock implies that they are not able to reduce the volatility of consumption as much as they wish. If given the choice, risk-averse agents would prefer to commit to a given sharing rule in order to smooth out future labor income shocks and only change it if there is a serious threat to the relationship.¹¹

3.2.8. Value of marriage. The spouses’ Bellman equations recursively characterize the equilibrium marital payoffs for given reservation values $(V_m^0(i), V_f^0(j))$. Consistently with the household problem and the divorce rule described in previous sections, the Bellman equations of married agents can be written as:

$$\begin{aligned}
V_m(l_f; i, j) = & v_m(l_f; i, j) + \psi_m(i)\beta \sum_{i'} V_m^0(i')\pi_m(i, i') + \\
& \underbrace{\psi_m(i)\psi_f(j)\frac{\beta}{2} \sum_{i', j'} [\gamma_e + \log(\exp(-\kappa) + \exp(\bar{S}(i', j')))] \pi_m(i, i')\pi_f(j, j')}_{\text{Husband's continuation value}}.
\end{aligned} \tag{3.2.15}$$

$$\begin{aligned}
V_f(l_f; i, j) = & v_f(l_f; i, j) + \psi_f(j)\beta \sum_{j'} V_f^0(j')\pi_f(j, j') + \\
& \underbrace{\psi_m(i)\psi_f(j)\frac{\beta}{2} \sum_{i', j'} [\gamma_e + \log(\exp(-\kappa) + \exp(\bar{S}(i', j')/\sigma))] \pi_m(i, i')\pi_f(j, j')}_{\text{Wife's continuation value}}.
\end{aligned} \tag{3.2.16}$$

where γ_e is Euler’s constant.

Enforcing the splitting rule (3.2.10) on the lhs of equations (3.2.15) and (3.2.16), one can recover the per-period indirect utilities $v_m(l_f; i, j)$ and $v_f(l_f; i, j)$. The pair $(v_m(l_f; i, j), v_f(l_f; i, j))$ needs to be such that the household allocation lies on the Pareto frontier: hence, substituting v_m and v_f in equation (3.2.6) ensures that the wife’s and husband’s payoffs are both feasible and Pareto optimal. Moreover, this also yields a Bellman equation for the

¹¹See [Ligon, Thomas, and Worrall \(2002\)](#) and [Mazzocco \(2007\)](#) for an exhaustive analysis of limited commitment: these papers discuss the ex-ante efficient contract. More recently, several other papers estimated intertemporal collective models with limited commitment: [Voena \(2015\)](#), [Reynoso \(2017\)](#) and [Shephard \(2018\)](#) are three noteworthy examples; see [Chiappori and Mazzocco \(2017\)](#) for a complete review.

surplus function:

$$\begin{aligned}
S(l_f; i, j) = & h(l_f, i, j) - \log \left[\exp \left(V_f^0(j) - \psi_f(j)\beta \sum_{j'} V_f^0(j')\pi_f(j, j') \right)^2 + \right. \\
& \left. + \exp \left(V_m^0(i) - \psi_m(j)\beta \sum_{i'} V_m^0(i')\pi_f(i, i') \right)^2 \right] + \\
& + \underbrace{\psi_m(i)\psi_f(j)\beta \sum_{i', j'} [\gamma_e + \log(\exp(-\kappa) + \exp(\bar{S}(i', j')))] \pi_m(i, i')\pi_f(j, j')}_{\text{Couple's continuation value}}
\end{aligned} \tag{3.2.17}$$

which yields a system of $|\mathcal{I}| \times |\mathcal{J}| \times |\mathcal{L}|$ equations with as many unknowns, the elements of $(S(l_f; i, j))$.¹² The function $h(l_f; i, j)$ depends on the shape of the Pareto frontier - their relationship is clarified in appendix 3.A.1 - and, ultimately, on preferences: in this case, it corresponds to the following

$$h(l_f; i, j) = \tilde{\gamma}_1(i, j) + (\gamma_1(i, j) + 1) \log(w_i + l_f w_j) + \gamma_2(l_f; j) + \gamma_3(i, j). \tag{3.2.18}$$

where $\tilde{\gamma}(i, j)$ is a function of $\gamma_1(i, j)$ (see appendix 3.A.1). By establishing the connection between preferences and the match payoff, the function h has a key role in the determination of the equilibrium marital patterns.

3.2.9. Meetings. Search for a partner is costless and singles of different sex meet each other randomly. In each period, the number of meetings between singles of different sex of a given type is given by a *meeting function* $\Lambda(i, j, n_{m,+}, n_{f,+})$. The measures $n_{m,+}$ and $n_{f,+}$ represent the available number of singles by type at time t_+ and are defined over $|\mathcal{I}|$ and $|\mathcal{J}|$ respectively: they will be formally defined in section 3.2.11, and are endogenously determined at equilibrium. The probability for a single man i of meeting a single woman j and the probability of a single woman j to meet a single man i can be written as the following conditional probabilities:

$$\Lambda_m(i, j) \equiv \Lambda(i, j, n_{m,+}, n_{f,+})/n_{m,+}(i) \tag{3.2.19}$$

$$\Lambda_f(i, j) \equiv \Lambda(i, j, n_{m,+}, n_{f,+})/n_{f,+}(j). \tag{3.2.20}$$

¹²A convenient way of proceeding is actually to derive the corresponding Bellman equation for \bar{S} by using the expression for the inclusive surplus (3.2.13). For given reservation values (V_m^0, V_f^0) , it is possible to solve the system by first computing \bar{S} for couples where at least one spouse has age \bar{a}_g : these couples have a continuation value equal to zero. Then, it is possible to solve backwards by computing \bar{S} for couples where at least one spouse has age $\bar{a}_g - 1$, and so on.

The meeting function $\Lambda(i, j, n_{m,+}, n_{f,+})$ needs to respect some theoretical restrictions. The total number of meetings involving types i (or j) cannot exceed $n_{m,+}(i)$ (or $n_{f,+}(j)$)¹³:

$$\sum_j \Lambda(i, j, n_{m,+}, n_{f,+}) \leq n_{m,+}(i) \quad \forall i \quad (3.2.21)$$

$$\sum_i \Lambda(i, j, n_{m,+}, n_{f,+}) \leq n_{f,+}(j) \quad \forall j \quad (3.2.22)$$

The specification chosen for Λ is the following:

$$\Lambda(i, j) = \lambda(i, j) \frac{n_{m,+}(i)n_{f,+}(j)}{n_{m,+}(i) + n_{f,+}(j)} \quad (3.2.23)$$

with $\sum_i \lambda(i, j) \leq 1$ for each j and $\sum_j \lambda(i, j) \leq 1$ for each i . These conditions on λ ensure that constraints (3.2.21) and (3.2.22) are respected. In line with the search literature, the number of meetings depends on the availability of singles on each side of the market (Rogerson, Shimer, and Wright, 2005). However, with heterogeneity on both sides of the market, the number of meetings between types i and j depend on the specific supplies $n_{m,+}(i)$ and $n_{f,+}(j)$: the two act as inputs in a CES function where the elasticity of substitution is equal to 1.¹⁴ Moreover, $\lambda(i, j)$ acts as a shifter that captures the degree of homophily along observable traits in the meeting structure: the empirical specification of $\lambda(i, j)$ is detailed in section 3.5.6.

3.2.10. Marriage decisions and value of singlehood. Upon a meeting, a man and a woman observe each other's type and draw a vector $\boldsymbol{\eta}$. As in equation (3.2.14), the logit framework yields the following probability of getting married:

$$\begin{aligned} \alpha_0(x, y) &\equiv \Pr \left\{ \max_{l_f \in \mathcal{L}} (S(l_f; i, j) + \eta_{l_f}) > 0 \right\} \\ &= 1 - \frac{1}{1 + \exp(\sigma_\ell \bar{S}(i, j))} \end{aligned} \quad (3.2.24)$$

where α_0 differ from α because of the absence of the sunk cost κ . In other words, if an agent is not interested in pursuing a relationship with his/her date, he/she needs to wait until the next period, but can walk away without having to pay any additional cost. As a consequence, it is easy to show that $\alpha(i, j) > \alpha_0(i, j)$ as long as $\kappa > 0$.

It is now possible to derive the value of singlehood, named V_m^0 and V_f^0 for men and women respectively. The per-period utility flow a single agent gets can easily be derived from a much simplified version of the household problem discussed for married couples. The only consumption choice a single agent needs to make is how to spend his total wage on

¹³In addition, at the aggregate level, the total number of meetings must not exceed $\min\{N_m, N_f\}$. This additional restriction is implied by (3.2.21) and (3.2.22) as long as there is an equal number of male and female singles on the market.

¹⁴The specification (3.2.23) generalizes $\Lambda(i, j) = \lambda(i, j) ((n_{m,+}(i))^{-\chi} + (n_{f,+}(j))^{-\chi})^{-1/\chi}$, widely used in demography (Pollak, 1990). When $\chi = 1$, Λ boils down to (3.2.23).

goods q and Q : since he/she lives alone, both goods are private. The presented discounted value of being single also incorporates the expectations about his/her marriage market prospects. The Bellman equations can be written as follows:

$$\begin{aligned}
V_m^0(i) = & v_m^0(i) + \beta \psi_m(i) \sum_{i'} V_m^0(i') \pi_m(i, i') di' + \\
& + \underbrace{\psi_m(i) \frac{\beta}{2} \sum_{i', j'} [\gamma_e + \log(1 + \exp(\bar{S}(i', j')))] \Lambda_m(i', j') \pi_m(i, i')}_{\text{Expected marriage prospects}} \quad (3.2.25)
\end{aligned}$$

$$\begin{aligned}
V_f^0(j) = & v_f^0(j) + \beta \psi_f(j) \sum_{j'} V_f^0(j') \pi_f(j, j') + \\
& + \underbrace{\psi_f(j) \frac{\beta}{2} \sum_{i', j'} [\gamma_e + \log(1 + \exp(\bar{S}(i', j')))] \Lambda_f(i', j') \pi_f(j, j')}_{\text{Expected marriage prospects}}. \quad (3.2.26)
\end{aligned}$$

3.2.11. Aggregate stocks. After outlining the behavior of individual agents, it is useful to define aggregate measures to keep track of the number of individuals by type and marital status in the population. First, consider the overall population dynamics: define p_m over \mathcal{I} , p_f over \mathcal{J} the marginal PDF of the male and female population characteristics. As detailed in section 3.2.1, agents may exit the marriage market in any period: e.g., the per-period aggregate outflow of men of type i is given by $(1 - \psi_m(i))p_m(i)$. Stationarity demands that the outflow of agents is counterbalanced by an inflow of new agents so that the size and composition of the population do not change. In t_+ in each period, an inflow $\omega_m(i)$ of unmarried men i and $\omega_f(j)$ of unmarried women j enter the market so that p_m and p_f do not change over time.¹⁵

I introduce measures n_m over \mathcal{I} , n_f over \mathcal{J} and m over $\mathcal{I} \times \mathcal{J}$ that count male singles, female singles and couples *at the end of the period*, i.e., after the matching phase took place (or, analogously, in t on time-lines 3.1 and 3.2, just before uncertainty is resolved). As in any matching model, the *matching outcome* (n_m, n_f, m) must respect the accounting restrictions:

$$p_m(i) = n_m(i) + \sum_j m(i, j) \quad (3.2.27)$$

$$p_f(j) = n_f(j) + \sum_i m(i, j). \quad (3.2.28)$$

¹⁵Relaxing this assumption introduces insightful long-run dynamics. For instance, younger cohorts may enter the market with better initial wages: in a competitive environment, this may have sizable implications for older cohorts' matching behavior. While such framework does not seem compatible with a notion of deterministic steady-state equilibrium, it may be helpful to study the relationship between business cycles and marriage markets.

In t_+ , when agents update their types, it is also necessary to update the aggregate distributions. Singles draw their new wages before the “market opening”, and they are joined by the inflows of new agents ω_m and ω_f . Hence, the measure of singles of a given type in t_+ is given by:

$$n_{m,+}(i') \equiv \begin{cases} \omega_m(i') + \sum_{i'} \psi_m(i) n_m(i) \pi_m(i, i') & \text{if } a_{i'} > \underline{a}_m \\ \omega_m(i') & \text{if } a_{i'} = \underline{a}_m \end{cases} \quad (3.2.29)$$

$$n_{f,+}(j') \equiv \begin{cases} \omega_f(j') + \sum_{j'} \psi_f(j) n_m(j) \pi_f(j, j') & \text{if } a_{j'} > \underline{a}_f \\ \omega_f(j') & \text{if } a_{j'} = \underline{a}_f \end{cases} \quad (3.2.30)$$

Similarly, it is useful to define a measure m_+ that counts married couples of a given type in t_+ , right after uncertainty is resolved, but before spouses could make decisions about the continuation of the match. In other words, the measure m_+ is the distribution of characteristics of the population at risk of divorce in t_+ :

$$m_+(i', j') = \sum_{i,j} \psi_m(i) \psi_f(j) m(i, j) \pi_m(i, i') \pi_f(j, j') di dj. \quad (3.2.31)$$

3.2.12. Law of motion and search equilibrium. The matching outcome (m, n_m, n_f) results from individual matching strategies (α, α_0) . The number $m(i, j)$ of couples (i, j) at the end of the period is given by the sum of newlyweds (i, j) and those couples (i, j) that did not divorce after drawing new wages and home productivity shocks. This results in the following law of motion:

$$m(i, j) = \underbrace{\alpha_0(i, j) \Lambda(i, j)}_{MF(i,j)} + \underbrace{\alpha(i, j) m_+(i, j)}_{m_+(i,j) - DF(i,j)}. \quad (3.2.32)$$

where the first term on the rhs provides a formula for the marriage flow $MF(i, j)$, while the second term implicitly provides a formula for the divorce flow $DF(i, j)$. Introducing the notation $NF(i, j) \equiv MF(i, j) - DF(i, j)$ to define the *net flow* of agents (i, j) , the evolution of stocks can also be described concisely as follows:

$$m(i, j) = m_+(i, j) + NF(i, j). \quad (3.2.33)$$

At the steady-state search equilibrium, agents' matching strategies (α, α_0) must be consistent with the equilibrium payoff structure S . The gains from marriage at equilibrium, described by S depend on both the household technology and the agents' reservation utilities (V_m^0, V_f^0) . The latter are endogenous equilibrium objects, in that they depend on the supplies (n_m, n_f) of singles on the market. Given these premises, the steady-state search equilibrium can be defined by combining the key equations outlined in this section.

DEFINITION 3. Consider the search-and-matching model described in this section. A **steady-state search equilibrium** is given by time-invariant measures of couples and singles (m, n_m, n_f) , payoffs (V_m^0, V_f^0, S) and strategies (α_0, α) so that:

- optimal marriage and divorce strategies (α_0, α) are linked with surplus S through the divorce rule (3.2.14) and the marriage rule (3.2.24).
- the Bellman equation of marital surplus (3.2.17) yields S for given reservation utilities (V_m^0, V_f^0) ;
- the Bellman equations of reservation utilities (3.2.25) and (3.2.26) yield (V_m^0, V_f^0) for given supplies of singles (n_m, n_f) ;
- if the accounting constraints (3.2.27) and (3.2.28) are enforced, the law of motion (3.2.32) yields the equilibrium aggregate measures (m, n_m, n_f) for given matching strategies (α_0, α) .

These equilibrium conditions can be combined to derive a fixed-point operator of type $n = T_{ext}n$ over the support $\mathcal{I} \cup \mathcal{J}$. This fixed-point operator is described in appendix 3.A.3. Recently, [Manea \(2017\)](#) generalized the original proof of existence by [Shimer and Smith \(2000\)](#). It may be possible to extend this proof to a framework with random match quality, although this has not been done in the literature yet. In practice, iteration of the fixed-point operator seems to converge to the same distribution n .¹⁶

3.2.13. Welfare measures. Solving for the steady-state search equilibrium yields the distribution of welfare across types. Following [Fernández and Wong \(2017\)](#), I focus on the welfare of individuals entering the marriage market at age \underline{a} . This corresponds to their reservation utility at age \underline{a} plus the expected surplus from participating to the marriage market for the first time (after drawing their first wage). In the case of young men, the *ex-ante welfare* can be computed as follows:

$$V_m^{ex}(i) = V_m^0(i) + \frac{1}{2} \sum_j [\gamma_e + \log(1 + \exp(\bar{S}(i, j)))] \Lambda_m(i, j). \quad (3.2.34)$$

The ex-ante welfare (3.2.34) accounts for both the *human wealth* of individuals and their expected gains from marriage over the life-cycle. It can be decomposed into two terms:

$$V_m^{ex}(i) = V_m^{hw}(i) + V_m^{mg}(i) \quad (3.2.35)$$

$$V_m^{hw}(i) = \mathbb{E}_{\mathbf{w}} \left[\sum_{a=\underline{a}}^{\bar{a}} \beta^{a-\underline{a}} v_m^0(h, a, w) \right]. \quad (3.2.36)$$

¹⁶The only caveat is that the introduction of a “relaxation parameter” is needed: details are provided in appendix 3.A.3.

The term $V_m^{hw}(i)$ accounts for the expected present discounted utility generated outside of the marriage market. In other words, if the agent were to stay single for his whole life, his expected present discounted utility would correspond to $V_m^{hw}(i)$. This term only depends on the agent’s human wealth, i.e., on the expected stream of labor market earnings obtained during his life-cycle. The term $V_m^{mg}(i)$ is defined as a residual in (3.2.35) and accounts for the additional expected present discounted utility generated through the marriage market. The term $V_m^{mg}(i)$ depends on the agent’s expectations about the timing and duration of his matches and about the characteristics of the partner(s).

The distribution of V_m^{ex} and V_f^{ex} provides insights on the equality of opportunities among young individuals and on the role of marriage markets in explaining welfare differences across groups and over time. [Fernández and Wong \(2017\)](#) use analogous measures to understand the implications of different divorce regimes across gender and wage groups. [Chiappori, Salanié, and Weiss \(2017\)](#) and [Reynoso \(2017\)](#) measure the *marriage market returns to education* by computing the difference in expected gains from marriage between those with a college degree and those without.

These measures are used later in section 3.6.9 in order to understand how the intertemporal welfare of young individuals reacts to changes in the primitive parameters of the model. In particular, changes in the wage structure affect agents’ human wealth and may increase or decrease the term $V_m^{hw}(i)$. However, they also affect the incentives to get married and the expected gains from marriage over the life-cycle. Changes in $V_m^{mg}(i)$ do not need to go in the same direction: they may either offset or amplify those welfare changes that occur outside of the marriage market.

As a final remark, the intertemporal measures of welfare used in the present draft do not allow me to assess the monetary costs incurred by households due to changes in the gains from marriage. One way to circumvent this problem is to employ a money metric measure of individual utility that is able to capture the monetary value of the gains from marriage ([Chiappori and Meghir, 2014](#)). The key idea behind their approach is that it is possible to provide married agents with a monetary compensation paid in case of divorce that makes them indifferent between keeping the current match and leaving: the said amount represents the monetary value of marriage. I plan to extend their methodology to the current dynamic framework in future work.

3.3. Identification

In this section, I discuss the identification of three key objects in the model, the meeting function Λ , the divorce cost κ , and the per-period match surplus h . I informally refer to the “full identification” of the model as the desirable situation where the observed data patterns can only be generated by a unique choice of (Λ, κ, h) , and where the primitive

parameters of the model can thus be inferred with appropriate data. I show that full identification can be obtained starting from a dataset $(\hat{n}_m, \hat{n}_f, \hat{m}, \hat{\ell}, \widehat{MF}, \widehat{DF})$, where $(\hat{n}_m, \hat{n}_f, \hat{m})$ is the observed marriage market outcome (the “stocks”), $\hat{\ell}$ is the observed vector of labor supply choices, $(\widehat{MF}, \widehat{DF})$ is the observed marital turnover (the gross “flows”).

3.3.1. Matching strategies. The identification of the match surplus starting from matched data has been exhaustively discussed by [Choo and Siow \(2006\)](#) and [Galichon and Salanié \(2015\)](#). [Choo \(2015\)](#) extends the seminal model by [Choo and Siow \(2006\)](#) to a dynamic framework where people age¹⁷: also in this case, identification of the gains from marriage relies on the observation of repeated cross-sections of matched data.

Similar identification principles apply to search-and-matching models: information on “stocks”, i.e., the number of married and single individuals by type, are still key to achieve the full identification of the model. However, the econometrician needs to address an additional question: are matches between two specific types i and j common (rare) because of a high (low) match surplus or because these types meet with high (low) frequency?

In this section, start by assuming that the meeting function Λ and the divorce cost κ are known to the econometrician. If this is the case, then it is possible to pin down matching strategies (α, α_0) using matched data $(\hat{n}_m, \hat{n}_f, \hat{m})$, starting from the law of motion (3.2.32) and the following relationship between α and α_0 :

$$\alpha_0(i, j) = \frac{\alpha(i, j)}{\tilde{\kappa} + (1 - \tilde{\kappa})\alpha(i, j)}. \quad (3.3.1)$$

where $\tilde{\kappa} = \exp(\kappa)$. The following identification result for $(\alpha(i, j), \alpha_0(i, j))$ applies.

LEMMA 2. *Denote θ the set of search parameters of the model (Λ, κ) and assume the econometrician observes $(\hat{n}_m, \hat{n}_f, \hat{m})$, where the empirical measures of singles \hat{n}_m and \hat{n}_f respect the empirical counterparts of the accounting constraints (3.2.27) and (3.2.28). Assume that, for a choice of θ and for each pair (i, j) , the following condition holds:*

$$\hat{m}(i, j) < \Lambda(i, j, \hat{n}_{m,+}(i), \hat{n}_{f,+}(j)) + \hat{m}_+(i, j). \quad (3.3.2)$$

Then there exist two unique mappings from the empirical distribution of spouses characteristics \hat{m} defined over $\mathbb{R}_+^{|\mathcal{I}| \times |\mathcal{J}|}$ to the matching strategies $\hat{\alpha}^\theta$ and $\hat{\alpha}_0^\theta$, both in $[0, 1]^{|\mathcal{I}| \times |\mathcal{J}|}$.

PROOF. Start from the law of motion (3.2.32), and substitute out $\alpha_0(i, j)$ with (3.3.1). For each (i, j) , $\hat{\alpha}^\theta$ is the one and only solution in the interval $[0, 1]$ to the quadratic

¹⁷While the frictionless model of [Choo \(2015\)](#) presents interesting similarities with the model outlined in this paper, one of the main differences is that in Choo’s paper marriage contracts are characterized by full commitment and the risk of divorce is fully exogenous.

equation:

$$\begin{aligned}
& -(\tilde{\kappa} - 1)\hat{m}_+(i, j)\alpha(i, j)^2 + \\
& + [(\tilde{\kappa} - 1)\hat{m}(i, j) + \Lambda(i, j, \hat{n}_{m,+}(i), \hat{n}_{f,+}(j)) + \tilde{\kappa}\hat{m}_+(i, j)] \alpha(i, j) + \\
& - \hat{m}(i, j)\tilde{\kappa} = 0.
\end{aligned} \tag{3.3.3}$$

Consider the generic notation for the quadratic equation $ax^2 + bx + c = 0$. Note that: $a < 0$ and $c < 0$ as $\tilde{\kappa} > 1$ by assumption; $\Delta \equiv b^2 - 4ac > 0$; the axis of symmetry is given by $x = -b/2a > 1$. It follows that both solutions take positive values, and that there is at most one solution between $[0, 1]$. The latter is between $[0, 1]$ if $2a < -b + \sqrt{\Delta}$: squaring both sides of the inequality yields $a + b + c > 0$, which holds by assumption (3.3.2). As long as $\hat{\alpha}^\theta(i, j) \in [0, 1]$, it is easy to see from the relationship (3.3.1) that also $\hat{\alpha}_0^\theta(i, j) \in [0, 1]$ regardless of the value of $\tilde{\kappa}$. \square

Condition (3.3.2) implies that, if Λ is misspecified, then there may not exist a set of individual strategies (α, α_0) that rationalize the observed matching outcome $(\hat{n}_m, \hat{n}_f, \hat{m})$. The rhs of condition (3.3.2) stands for the total number of pairs (i, j) that, in t_+ , are considering whether they should spend the next period together. If the number of meetings (i, j) is too low, there may not be a sufficient number of potential matches to rationalize the net flow $NF(i, j)$. In other words, underrating the number of meetings (i, j) will lead $(\hat{\alpha}, \hat{\alpha}_0)$ to exceed their upper bounds.¹⁸ The set of conditions (3.3.2) can thus be used to test whether the prior on the specification chosen for Λ is to reject.

However, we are now left with a major question to address. While conditions (3.3.2) imply some restrictions on the meeting structure, there may still exist several functions Λ that satisfy them. All of these meeting functions would be able to rationalize the observed matching outcome $(\hat{n}_m, \hat{n}_f, \hat{m})$, with the econometrician being unable to distinguish among them. The intuition behind this identification puzzle is the following: a cross-section (or a series of cross-sections) of matched data are only consistent with a unique set of net flows \widehat{NF} , as it is clear from the law of motion (3.2.33). However, there may be several sets of gross flows $(\widehat{MF}, \widehat{DF})$ consistent with the same dataset. This identification issue is addressed in the next section.

3.3.2. Meeting and marriage cost function. GJR formally discuss this identification problem in a search-and-matching framework without aging and wage shocks. They suggest the use of additional data on gross flows in order to disentangle the structure of

¹⁸Interestingly, condition (3.3.2) does not imply any restriction on the cost structure, apart from $\kappa > 0$. The proof can be easily extended to the case where the covariance structure of the initial shock differs from the covariance of the following shocks: also in this case condition (3.3.2) does not depend on either κ or the covariance of the first shock. However, a proper generalization of this result would require to relax the GEV assumption on the home productivity shocks, which would result in a generalization of relationship (3.3.1). Once relaxed the distributional assumption, one may end up facing restrictions on both Λ and κ .

meetings from the structure of the surplus. Assume now that the econometrician observes a new layer of data $(\widehat{MF}, \widehat{DF})$. Recall from the law of motion (3.2.32) that the equations that yield marriage and divorce rates are:

$$\alpha_0(i, j) = \frac{MF(i, j)}{\Lambda(i, j)} \quad (3.3.4)$$

$$1 - \alpha(i, j) = \frac{DF(i, j)}{m_+(i, j)}. \quad (3.3.5)$$

Relationships (3.3.4) and (3.3.5) imply $2 \times |\mathcal{I}| \times |\mathcal{J}|$ restrictions that can be used to achieve full identification of the model. Hence, the number of unknowns in the system implied by (3.3.4) and (3.3.5) cannot exceed $2 \times |\mathcal{I}| \times |\mathcal{J}|$. The first $|\mathcal{I}| \times |\mathcal{J}|$ parameters to estimate are the matching probabilities α for each pair (i, j) . The only additional unknown parameter needed to obtain the marriage probabilities α_0 through relationship (3.3.1) is the cost of divorce κ . Hence, it is necessary to impose a restriction on Λ as it is only possible to identify up to $|\mathcal{I}| \times |\mathcal{J}| - 1$ additional parameters.

The key intuition behind this identification strategy is that both marriage and divorce flows contain information about marital surplus. The structural approach helps establish an explicit relationship between the data and the parameter of interest. Similar empirical strategies have already been used in the matching literature. In the case of marriage markets, [Wong \(2003\)](#), [Goussé \(2014\)](#) and GJR target the rate of arrival of meetings and divorce shocks: the first two papers rely on data on the duration of marriage and singlehood, while the third relies on marriage and divorce rates. In this paper, I use an approach analogous to GJR, although I also exploit the variation in marriage and divorce rates across types to estimate a more general meeting function. Interestingly, [Bruze, Svarer, and Weiss \(2015\)](#) use a similar strategy to estimate a frictionless model of the marriage market with heterogeneous cost of divorce in terms of both individual types and duration: they exploit the variation in the hazard of divorce with respect to marriage duration. Finally, both [Greenwood, Guner, Kocharkov, and Santos \(2016\)](#) and [Shephard \(2018\)](#) assume that the unobserved match-quality is autocorrelated over time: information on marriage duration and divorce can also help with the identification of the parameters of the match-quality distribution.

3.3.3. Surplus and household production function. Using the identification result from last section, it is possible to infer the cost of divorce and the meeting function: their estimates are named $\hat{\theta} \equiv (\hat{\Lambda}, \hat{\theta})$. Hence, it is possible to back out the matching strategies α and α_0 from the law of motion (3.2.32), as long as the conditions required by lemma 2 are respected. The average surplus \bar{S} can be obtained through the bijection (3.2.14). Data on labor supply - and in particular knowing the proportion $\hat{\ell}(l_f; i, j)$ of couples (i, j) with l_f hours worked in the data - leads to the identification of the surplus function S over the support $|\mathcal{I}| \times |\mathcal{J}| \times |\mathcal{L}|$ through equation (3.2.12). The identification

of the reservation utilities V_m^0 and V_f^0 follows, provided a normalization of the utility flows for singles, $v_m^0(i)$ and $v_f^0(j)$.¹⁹

Another important remark is that, as long as the wage process (π_m, π_f) and the survival rates (ψ_m, ψ_f) are independent of the marital status, they can be estimated outside of the model. This greatly simplifies the estimation, but overcoming this limitation is a necessary step to fully understand the relationship between the changing wage distribution and marriage market outcomes. At the moment, only [Shephard \(2018\)](#) and [Beauchamp, Sanzenbacher, Seitz, and Skira \(2018\)](#) have been successful in estimating a model of marriage market equilibrium with human capital accumulation during the life-cycle, while several papers have explored these issues outside of the marriage market equilibrium (e.g. [Mazzocco, 2007](#); [Blundell, Dias, Meghir, and Shaw, 2016](#)).

Under these assumptions, it is possible to pin down the per-period match surplus h from the Bellman equation for surplus (3.2.17), as by now both the surplus and the reservation values are known for given search parameters $\hat{\theta}$. In practice, h is obtained as the residual after subtracting both the current reservation values and the dynamic component (i.e., the continuation value) from the surplus function S .

3.4. Data

The estimation closely relies on the identification results derived in the previous section. Hence, in order to identify the key primitives of the model - the meeting and the per-period match surplus - two types of data are needed: (i) the standard *matched data*, i.e., cross-sectional data on who is matched with whom, and (ii) panel data to measure the *hazard of marriage* (among singles) and the *hazard of divorce* (among married). In this paper, I use two separate data sources to obtain all the necessary information: the Annual Social and Economic Supplement (ASEC) from the Current Population Survey (CPS) is used as a source of information on the number of singles and married couples by observable characteristics, while the Panel Study of Income Dynamics (PSID) as a source of information on the hazard of marriage and divorce. In this section, I introduce the sample used for estimation and the definition of the main variables; I clarify how I use the two datasets and describe their salient characteristics.

3.4.1. Sample selection. The CPS is composed of a series of yearly cross-sections, and observations are assigned individual cross-sectional weights so that the sample is representative of the American population in a given year. From the main CPS dataset, I

¹⁹In the literature on the econometrics of matching models, a normalization of the agents' reservation utilities is usually required as matched data only identify the match gains, i.e., the differential utility produced by the match ([Galichon and Salanié, 2015](#)). In search models, an analogous normalization applies to the per-period utility flows of singles.

build two separate samples: in the first, I pool all individuals aged between 20 and 60 observed between 1971 and 1981 (about 409,000 men and 445,000 women); in the second, I pool all individuals in the same age range observed between 2001 and 2011 (about 613,000 men and 664,000 women).

From the main PSID dataset, I only keep observations from the SEO (Survey of Economic Opportunity) and the SRC (Survey Research Center) samples, thus excluding the Immigrant and Latino samples. Since 1997, the survey has been conducted every two years: since in this analysis the PSID is only used to exploit its panel dimension, I only keep odd years in the sample and focus on changes that occur over a two-year period. Also in the case of the PSID, I build two separate samples: the first contains all individuals aged between 20 and 60 between 1971 and 1981 (about 9,800 observations for men and 11,100 for women), and the second all individuals in the same range between 2001 and 2011 (about 12,200 observations for men and 13,300 for women). In the PSID, attrition is low for sampled individuals: this yields a fairly balanced panel. However, temporary non-sample individuals living with the sampled are not followed once they quit the household: this matters when looking at the distribution of divorcees, as it is not always possible to follow the trajectory of one of the two partners after the breakup.

In both samples, I rely on the “relationship-to-head” variable in order to identify married couples. An important remark is in order: the CPS only surveys the civilian noninstitutional population, thus excluding people on duty in the Armed Forces or living in correctional institutions or long-term care facilities. This may explain why the number of observations is higher for women; once statistical weights are implemented, the women-to-men gender ratio is about 0.51, both in the 1970s and in the 2000s.

3.4.2. Main variables.

- *Conjugal status*: both PSID and CPS respondents are associated with a household identifier and a “relationship-to-head” variable. I identify couples through the presence of an individual who claims to be the head’s legal spouse. If a person is head of household and not living with a spouse, I consider him/her to be single. In addition, if he or she is living in a household and is not the head, nor partner of the head, I also consider him/her to be single. As anticipated, members of secondary families (e.g., a child of the head living with his/her partner) are excluded from the sample. In the CPS, information about unmarried cohabiting partners are only available starting from 1995: hence, the empirical analysis focuses on married couples only, and individuals in cohabiting couples are counted as singles.²⁰

²⁰Since information on unmarried partners has been available since the earliest wave of the PSID, I plan to conduct robustness checks with a sample taken for the 2000s period where I consider all couples - married and cohabiting - as equal. [Lundberg and Pollak \(2014\)](#) stress that, once cohabitation is taken

- *Education*: I divide respondents into two categories, those with a college degree and those without. Given the cross-sectional nature of the CPS, education is taken to be the highest diploma achieved at the moment of the survey. Because of this, On the contrary, this problem can be avoided in the PSID: I define education as the highest level achieved along the longitudinal dimension.
- *Hours worked*: in both the CPS and the PSID, the average number of hours worked per week is defined as the total number of hours worked in a year divided by the number of weeks worked in a year. I then build a discrete labor supply variable as follows: agents working less than 7 hours per week on average are considered to be out of the labor force (NW); if they work between 7 and 34 hours, I define them as working part-time (PT); if they work 35 hours or more, I define them as working full-time (FT). The upper bound for part-time workers (34 hours) is consistent with the definition of the Bureau of Labor Statistics.²¹
- *Hourly wage*: in both the CPS and the PSID, hourly wages are obtained by dividing the (deflated) total yearly labor income by the total yearly number of hours worked. It is initially set to zero for those individuals who are out of the labor force. A drawback of the PSID dataset is that information on earnings and labor supply is only available for surveyed heads and partners, but is not available for other members of the household: hence, for some sampled individual-year observations, labor market information is missing, particularly in earlier waves when sampled individuals are young and more likely to live with their parents. The problem of selection due to labor force participation choices is addressed in the estimation phase.
- *Newlyweds*: changes in household composition are only observed in the PSID thanks to its longitudinal panel dimension. Changes in conjugal status occurring between two PSID waves help identify the formation of new couples. When a respondent is single in year $y - 2$ and living with a legal spouse in year y , he or she is defined as a “newlywed” in year y . Note that incoming spouses are typically not included in the PSID sample before the marriage occurs.
- *Divorcees*: in a similar way, it is possible to identify divorces in the PSID. If a married couple is observed living together in year y and at least one of the two partners (typically the sampled individual) is observed living alone or with a different spouse in $y + 2$, then the couple is flagged as “about to divorce”. I do not make a difference between divorce and separation as the legal duties of divorcees are not taken into account in the analysis. This way of identifying dissolving couples allows me to gather information on both spouses’ characteristics: as a comparison, datasets containing retrospective information on marital history do

into account, the gap between college and non-college graduates in the odds of being in a two-adult household is reduced.

²¹I have not tried to use a different definition of part-time yet, although I intend to do so in the future to check whether current findings are consistent.

not provide detailed information about former partners. In spite of this, since in some cases couples are composed of one followable and one non-followable spouse, it is not always possible to track the trajectories of both after a divorce.²²

3.5. Empirical specification and estimation

In this section, I provide details about the estimation method, which is composed of multiple steps. First, I estimate the age and education profiles of hourly wages and the parameters of the $AR(1)$ process for both the CPS and the PSID. The estimation of the wage distribution is performed out of the model. Second, I estimate the marriage surplus jointly with the search parameters (the meeting function and the cost of divorce). Last, starting from estimates obtained at the previous step, I recover the parameters of the production function of the public good (see function (3.2.2)).

3.5.1. Wage levels. In both the CPS and PSID, selection into the labor force prevents me from observing the wage distribution of the entire population. In order to address this issue, I take a control function approach and estimate a standard selection model à la Heckman (1974) for each gender and broad educational group (college graduates and non-college graduates). In the selection equation, I include the number of children, and the ages of the youngest and eldest child as instruments. I subsequently replace missing wages with predicted wages: these are used to assign individuals to wage quantiles, conditionally on their age and education group.

I use the wage distribution obtained with CPS data after estimating the selection model to compute the wage levels used in the following estimation steps. Each individual is assigned the median wage computed within his/her wage quantile, conditionally on his/her age and education. For instance, observations that rank in the top quintile of their age and education group are assigned a wage rate corresponding to the 90th percentile of the group.

It is important to remark that, to be consistent with the model of marriage market and labor supply outlined in previous sections, the wage distribution should be estimated jointly with the other parameters of the model. In this sense, the present work should only be regarded as a first step to extend GJR to include labor force participation choices. In addition, the instruments used to estimate the selection model are far from being ideal: Chiappori, Dias, and Meghir (forthcoming), who also opt for an estimation of the age profile of wages outside of the model, suggest the use of policy variation in out-of-work

²²Note that all individuals belonging to the original 1967 PSID sample are followable: hence, in the early waves of the PSID, the majority of couples is made of two followable individuals. However, couples that formed later (i.e., where an incoming spouse joined a sampled individual) are made of one followable and one non-followable individual.

income as an instrument for participation, an approach that could be explored in this framework. As a final remark, note that, in spite of these limitations, I discretize the wage support and only use a limited number of moments from the wage distribution: this should also limit the implications of misrepresenting such distribution.

3.5.2. Wage mobility. I have so far described the cross-sectional distribution of wages without imposing any restriction on wage mobility. In order to characterize the degree of mobility implied by the wage process, I follow [Bonhomme and Robin \(2009\)](#) and map the marginal distributions of w_i and $w_{i'}$ into the joint distribution of $(w_i, w_{i'})$ by using a copula. This convenient representation of the $AR(1)$ wage process leaves complete freedom on the way of specifying the marginal distribution of wages, described in the previous paragraph. Call $r_i(h_i, a_i)$ the rank of w_i among agents with human capital h_i and age a_i . The joint CDF of the current wage rank $r_i(h_i, a_i)$ and the future wage rank $r_{i'}(h_{i'}, a_{i'})$ is given by a Plackett's copula $C_m(r_i(h_i, a_i), r_{i'}(h_{i'}, a_{i'})|h_i, a_i)$ (with $C_f(r_j(h_j, a_j), r_{j'}(h_{j'}, a_{j'})|h_j, a_j)$ for women). A Plackett's copula is characterized by a single parameter that can be interpreted as a measure of mobility: this one parameter is a monotonically increasing function of the Spearman's rank correlation coefficient (details are provided in appendix 3.A.4). I allow this correlation coefficient to vary with agent's gender, education and age, and denote the vector of coefficients $\rho_m(h_i, a_i)$ and $\rho_f(h_j, a_j)$, for men and women respectively. To estimate these parameters, panel data on wages are needed: hence, after dealing with non-participation as explained in the previous section, I use PSID data to estimate the Spearman's coefficients $\rho_m(h_i, a_i)$ and $\rho_f(h_j, a_j)$ as detailed in [Bonhomme and Robin \(2009\)](#). The results are plotted in figure 3.5.

This setup implies that the degree of wage mobility faced by agents in the model is effectively summarized by $\rho_m(h_i, a_i)$ and $\rho_f(h_j, a_j)$: for a man of type i , the odds of moving up or down the wage distribution when getting older entirely depend on $\rho_m(h_i, a_i)$. In their analysis of earning mobility in the US, [Kopczuk, Saez, and Song \(2010\)](#) suggest that rank correlation between periods is indeed a direct and effective measure of wage mobility.

3.5.3. Agents' types and marginal distributions. The time-invariant human capital endowment, h_i for men and h_j for women, is assumed to correspond to education. I consider two levels of education, college graduates vs non-college graduates. Hence, $h_i, h_j \in \{1, 2\}$, where 2 stands for a college degree. All agents enter the marriage market at age 20, a period corresponds to two years, and all agents quit the market when aged 62: the life-cycle stretches across 21 periods, $a_i, a_j \in \{1, 2, \dots, 21\}$. Finally, for each age and education group, I rank agents according to their wage rates and divide them into wage quantiles: men (women) within the same wage quantile are *ex ante* identical, and are all assigned the same wage rate w_i (w_j), which corresponds to the median wage within

the quantile, consistently with what explained in section 3.5.1. I use 3 quantiles in the estimation phase, and 5 in the simulations and counterfactual exercises.

I assume that the cohort size is constant: hence, the marginal distribution with respect to age is uniform for both men and women. I also assume that the gender ratio is perfectly balanced. This assumes away gradual entry and exit into the marriage market, as well as the potential implications of gender ratio imbalances. While these restrictions could easily be relaxed, they allow me to simplify the analysis and to focus on a smaller number of primitive parameters in the empirical analysis. The only parameter left to estimate is the share of college graduates in the male and female population: I measure it as the proportion of college graduates older than 21 in the CPS. The share of adult college graduates per decade is given in table 3.3.

If the stationarity assumption were verified in the data, the population would be uniformly distributed with respect to age in the age bracket spanning from 20 to 60, once accounted for exogenous entry and exit.²³ This would be consistent with the assumptions stated in the previous paragraph. However, the stationarity assumption implies that both the size of the US population and its educational composition do not change over time. It is easy to see that these conditions are not verified in the data: in both the 2000s and the 1970s sample, younger individuals tend to be more educated than the older, as a consequence of the steady rise in college attendance. On top of that, in the 1970s sample, younger individuals are more numerous than the older: this is due to the exceptional demographic growth in the Post-war period. These demographic patterns would require to study the out-of-steady dynamics spurred by a growing and increasingly educated population: the changing size and composition of the population may have important implications in marital sorting across cohorts. This challenging issue is beyond the scope of the current paper, and is left for future search.

3.5.4. Marital patterns and hazard rates. The estimation of the model relies on the observation of the empirical frequencies of couples by type (i, j) , male singles by type i and female singles by type j . Hence, as a first step, I compute the raw empirical frequencies from the CPS sample, using the repartition into wage quantiles explained in section 3.5.1. These can be read as a contingency table - men's types on one axis and women's type on the other - and automatically imply some marginal frequencies through the accounting constraints (3.2.27) and (3.2.28). However, recall that, in section 3.5.3, I have already produced estimates of the marginal distributions \hat{p}_m and \hat{p}_f that are consistent with the stationarity assumption. To reconcile marital patterns and marginal distributions, I follow [Greenwood, Guner, Kocharkov, and Santos \(2014\)](#) and apply an

²³By tracking a cohort of agents born in the same year across CPS waves, it is possible to verify whether the cohort size is constant or is modified by inflows or outflows of people. For young agents, death rates are low; however, immigration, incarceration and active duty in the armed forces in the army do, among other factors, result in non-negligible variation in cohort size across years in the CPS.

iterative procedure to obtain what they define as “standardized contingency tables”, i.e., standardized empirical frequencies that are, at once, consistent with the required marginals \hat{p}_m and \hat{p}_f and with the matching behavior observed in the data.²⁴ The joint distribution of characteristics and marital status $(\hat{m}, \hat{n}_m, \hat{n}_f)$ obtained through this standardization technique is used to estimate the model in the next step.

The estimation also requires information on the hazard of marriage and divorce for different groups of the population. These are calculated from the PSID sample.

- Divorce rates for men of type i (women of type j) are calculated as the ratio of men i (women j) divorcing between t and $t + 1$ to the number of married men i (women j) in t . Similarly, for couples (i, j) , the rate corresponds to the ratio of the number of divorcing couples (i, j) between t and $t + 1$ to the number of married couples (i, j) in t .
- Marriage rates for men of type i (women of type j) are calculated as the ratio of men i (women j) getting married between t and $t + 1$ to the number of single men i (women j) in t . For couples (i, j) , the rate corresponds to the ratio of the number of couples (i, j) getting married between t and $t + 1$ to the geometric mean between the number of male singles i and female singles j in t .

In practice, I do not compute the full distribution but only some selected moments (e.g., the rate of divorce for men with a college degree). This allows me to avoid dealing with empty cells due to the small size of the PSID.

3.5.5. Search parameters and marriage surplus. The unobserved parameters of the meeting function Λ and the cost of divorce κ are estimated jointly with the vector of matching strategies α . The estimation procedure relies on matching a vector $\hat{\mu}$ of empirical moments calculated from the empirical distribution $(\hat{M}F, \hat{D}F)$ to their theoretical counterparts $\mu(\theta, \alpha^\theta)$. The vector $\hat{\mu}$ includes both rates of marriage and divorce conditionally on the spouses’ education and age.

In practice, I use an estimator of the GMM class and exploit the restrictions implied by the model on the size of the gross flows into and out of marriage. These correspond to equations (3.3.4) and (3.3.5), where the matching strategies α and α_0 are replaced by their nonparametric estimators, whose existence is guaranteed by Lemma 2. The GMM estimator is given by:

$$\hat{\theta} \equiv \arg \min_{\theta} \sum_k \omega_k (\mu_k(\theta, \alpha^\theta) - \hat{\mu}_k)^2 \quad \text{subject to (3.3.2) for any } (i, j) \quad (3.5.1)$$

²⁴More explicitly, this technique suggested by [Mosteller \(1968\)](#) consists in transforming a contingency table into a second one that respects the desired marginals while leaving the odds ratio unchanged. In the case of marriage, the transformation does not affect the odds of marrying a type i versus a type i' , for each woman j (and vice versa). Details on the computation are provided in appendix 3.A.5.

where ω is a vector of weights.²⁵ The presence of the constraint (3.3.2) is required by Lemma 2 to ensure that α^θ belongs to the set $[0, 1]$ and is indeed a probability. Since the number of constraints is large and corresponds to $|\mathcal{I}| \times |\mathcal{J}|$, I add a penalty function to the standard quadratic loss function used as objective. If the penalty function is given enough weight, this results in only small violations of the constraints.²⁶ Standard errors are computed following the guidance of [Lise and Robin \(2017\)](#).

Once obtained the estimates of the search parameters $\hat{\theta}$, I compute the corresponding matching strategies $\hat{\alpha}$ and the surplus function \hat{S} from the matching rule (3.2.24). Inverting equation (3.2.12), I then recover the surplus function \hat{S} - i.e., the match surplus specific to a given labor supply choice l_f - using the odds of choosing l_f for a married couple (i, j) observed in the data. Then, I derive the function h - which implicitly characterizes the Pareto frontier - from (3.2.17), the Bellman equation of surplus: this requires subtracting the continuation value and the reservation values terms, which are by now known terms, from the nonparametric estimate \hat{S} of the surplus function. The intuition is that, once taken into account the dynamic nature of surplus, the per-period match surplus $h(l_f; i, j)$ represents the residual component of the gains from marriage.

A remark is in order concerning the role of the parameter σ_ℓ - i.e., the degree of independence between domestic productivity shocks $\boldsymbol{\eta}$. This parameter weighs the importance of the economic component of the marital surplus, which ultimately depends on the labor supply choice of the wife, as opposed to that part of the marital surplus that only depends on the agents' traits (γ_3 in the per-period match surplus (3.2.18)). The parameter σ_ℓ is identified through the variation in labor supply behavior conditional on the agents' observable traits and sorting patterns along the same observables: its identification has been widely discussed in the nested logit literature ([Train, 2009](#)). However, I have not attempted to estimate σ_ℓ : so far, I have set $\sigma_\ell = 0.7$ to allow for some positive correlation between the shocks $\boldsymbol{\eta}$. There is little if no benchmark in the literature in this regard, although [Train \(2009\)](#) discusses its lower and upper bound ($0 < \sigma_\ell \leq 1$).²⁷

Finally, note that this estimation method exploits a set of restrictions produced by the model and has the advantage of being relatively quick to implement, as it is not necessary to solve for the market equilibrium. Moreover, the estimator is of the GMM class and has well-known properties. In [Ciscato \(2018\)](#), I employ a parallel Markov Chain Monte Carlo method to estimate the model in one step and without having to rely on the full distribution $(\hat{m}, \hat{n}_m, \hat{n}_f)$ but only using some of its moments. This approach is far more time-consuming

²⁵In the application, $\omega_k = 1$ for each k .

²⁶This means that α^θ may be slightly larger than 1. If the weight of the penalty function is high enough, violations are small. Once obtained $\hat{\alpha}$, for those types (i, j) for which the constraint is binding $\hat{\alpha}(i, j)$ is rounded down to 0.999.

²⁷While so far my findings have not proved to be particularly sensitive to changes in this parameter, I plan to conduct a more systematic analysis to show that they are indeed robust in this regard. Ultimately, I also plan to estimate σ_ℓ with the other unobserved parameters of the model.

as it requires to iteratively solve for the equilibrium model. However, it would allow me to estimate the parameters of the domestic production function, of the meeting function, the cost of divorce in one step. In addition, this kind of moment-matching approach does not require the *i.i.d.* assumption on the match-quality shocks.

3.5.6. Empirical specification: meeting function. The general specification (3.2.23) chosen for the meeting function Λ contains the unobserved parameters $\lambda(i, j)$, estimated jointly with the cost of divorce κ through the procedure described in section 3.5.5. I consider the following specification for the shifter $\lambda(i, j)$, which captures homophily in meetings:

$$\tilde{\lambda}(i, j) = \begin{cases} \frac{\exp(\lambda_1 \{h_i=h_j\} + \lambda_2 d^2 + \lambda_3 a_m a_f^2 + \lambda_4 a_m^2 a_f)}{1 + \exp(\lambda_1 \{h_i=h_j\} + \lambda_2 d^2 + \lambda_3 a_m a_f^2 + \lambda_4 a_m^2 a_f)} & \text{if } \underline{d} \leq d \leq \bar{d} \\ 0 & \text{otherwise.} \end{cases} \quad (3.5.2)$$

with $d \equiv a_m - a_f$. The terms of λ have the role of shifting the odds of meetings across types: the first term of (3.5.2) determines the degree of homophily with respect to education, while the last three let the odds of meeting depend on the age distance. On top of this, I impose that meetings do not occur at all if the age distance is too large: in the empirical analysis, I set $\underline{d} = -3$ and $\bar{d} = 7$, i.e., I do not allow for meetings between men that are more than 6 years younger than women or more than 14 years older than women (as one period corresponds to two years in the data). This restriction still allows me to consider about 96% of all marriages in the CPS for both the 1970s and the 2000s. The specification is close to the one suggested by [Shephard \(2018\)](#).

3.5.7. Empirical specification: household production function. With data on matching behavior and the labor supply of married couples, I have shown how to compute h on each point of its support $\mathcal{I} \times \mathcal{J} \times \mathcal{L}$. We have seen in section 3.2.8 that the specification of h depends on the agents' utility functions and the production function (3.2.2): equation (3.2.18) allows me to establish a connection between the parameters $(\gamma_1, \gamma_2, \gamma_3)$ and the per-period surplus h . Hence, through OLS, I compute estimates $(\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3)$ that best fit the vector of residuals \hat{h} .

I impose the following restriction on the parameter functions γ_1 , $\gamma_2(l_f)$ and γ_3 , which characterize the production of the public good Q through function (3.2.2).

- γ_1 , which represents the elasticity of Q with respect to the total joint expenditure $t_m + t_f$, only depends on the spouses' time-invariant human capital (h_i, h_j) . In particular, I estimate γ_1 for each combination (h_i, h_j) .
- $\gamma_2(l_f)$, which represents the productivity shifter associated with labor supply choice l_f , only depends on the wife's age a_j and human capital h_j . Hence, I

estimate γ_2 for each combination (h_j, a_j) and each level l_f , after normalizing $\gamma_2(1) = 0$, where $l_f = 1$ corresponds to the wife working full-time.

- γ_3 , which is an additional productivity shifter, only depends on the wife's age a_j and human capital h_j and on the husband's a_i and h_i , as well as interactions between these inputs. In particular, γ_3 contains dummies for each combination (h_i, a_i, h_j, a_j) . Hence, while γ_3 can change with age, its trajectory is fully predictable at the moment of the match, since it is not affected by wage shocks.

3.6. Results

In this section, I start by briefly discussing the estimates of the parameters of the domestic production function of the public good and the meeting function. I subsequently present the fit of the model for the two samples, the 1970s and the 2000s. I then conduct a counterfactual analysis to understand the main forces behind the decline of marriage that occurred between these two periods and discuss the changes in welfare. I conclude with a summary of the main findings and a concise description of the economic mechanisms at play.

3.6.1. Estimation results: public good. Estimates of the parameter γ_1 are collected in table 3.1, and suggest that not all families enjoy increasing returns to scale from the joint expenditure $t_m + t_f$ used to produce the public good Q . In the 1970s, the estimate of γ_1 for couples where both spouses are college graduates is not significantly different than 1; however, in the 2000s, these same couples enjoy significant increasing returns to scale. Conversely, couple where no spouse holds a college degree have experienced a drastic decrease: γ_1 used to be significantly greater than 1 in the 1970s, while it is not in the 2000s.

Estimates of the parameter γ_2 suggest that the gains from marriage for couples where the wife does not work are lower in the 2000s than in the 1970s, particularly for non-college graduates. For women aged between 40 and 50 in the 2000s, the benefits of staying at home are not significantly different than zero, regardless of the educational level. Part-time does not seem to produce public benefits, and, on the contrary, it may decrease the match surplus. It is possible to conclude that women have higher incentives to spend time on the labor market in the 2000s than they used to in the 1970s: the production of public good Q has become less labor-intensive.

3.6.2. Estimation results: meetings and cost of divorce. Estimates of the parameters of the meeting function Λ suggest that the marriage market in the 1970s is characterized by a higher segregation with respect to education than in the 2000s. The

odds of meetings across educational levels have increased between the two periods, all else constant: this is captured by the decrease in the parameter λ_1 reported in table 3.2.

The cost of divorce κ has increased by 0.4 standard between the two periods, although it is already high in the 1970s. Marriage is associated with strong commitment in both periods.²⁸ Recall that, in the model, it is assumed that spouses have the option to renegotiate the household allocation in every period: this assumption may cause the estimate of κ to be inflated. If the spouses were able to reduce consumption volatility by sticking to an agreed sharing rule, then they could achieve higher level of surplus as the continuation value of marriage would be higher. If this is the case, the source of commitment is misinterpreted and attributed to high costs of divorce.

3.6.3. Fit of the model. The estimated model is able to reproduce some of the key facts observed in the data, and does a good job at reproducing the changes in marital patterns that have occurred between the 1970s and the 2000s. Tables 3.4 and 3.5 compare simulated and empirical moments from the distribution of singles' and spouses' characteristics in the cross-section. Table 3.5 also contains the simulated and empirical hazard rates of marriage and divorce. For the 1970s sample, the predicted share of married men aged between 20 and 60 is 74.59%, as opposed to 69.74% in the data; for the 2000 sample, the predicted share is 58.64%, as opposed to 52.14% in the data. The distance between simulated and empirical moments is slightly larger for women. Most importantly, the estimated model replicates the decline of marriage between the 1970s and the 2000s accurately (see last column of table 3.4). The differences in the share of married adults between college graduates vs non-college graduates is also understated. However, the estimated model accurately replicates the differential in the decline of marriage across educational groups that has emerged between the 1970s and the 2000s. A graphical representation of how well the model is able to replicate the changes on the extensive margins is provided in figures 3.9 and 3.10.

The estimated model is able to approximate the age profile of married individuals: the simulated stock of married agents progressively increases during their 20s and 30s and stabilizes around age 40. The share of married women decreases when they enter their 50s: those who divorce or become widows outnumber those who get married, as women tend to marry older men. Note that, in the model, agents die when they turn 62, so this artificially inflates the number of widows. From figures 3.7 and 3.8, we can also see that the stock of married agents grows rapidly in the 1970s and more slowly in the 2000s.

The predicted correlation rates between spouses' ages and wages are only slightly below those computed from the data. The predicted educational homogamy is instead stronger.

²⁸Consider a potential couple (i, j) with the systematic part of surplus, $\bar{S}(i, j)$, equal to zero: their odds of getting married are one out of two. However, with $\kappa = 4.6$, if the man i and the woman j were already married, the probability that they will file divorce is only equal to about 1%.

The predicted changes between the 1970s and the 2000s are in line with what observed in the data. Finally, the aggregate labor force participation of women is well replicated by the model.

The predicted marriage and divorce rates are lower than those observed in the data (see table 3.5). This also affects the evolution of the stocks along the life-cycle (see 3.7 and 3.8): (i) agents are very selective, thus marriage rates are lower than in the data and the stocks of married agents grow slower; (ii) divorce is rare and the stocks of married agents keep increasing, albeit more slowly, instead of stabilizing. Low marriage and divorce rates are induced by the high point-estimates for the cost of divorce κ jointly with the relatively low R^2 obtained at the second stage of the estimation procedure (0.15 with the 1970s sample and 0.23 with the 2000s sample). In the future, estimating all parameters in one step - and thus fitting both the stocks and the flows at the same step - may address this issue.

3.6.4. Explaining the decline of marriage: full decomposition. In order to address the key question of this paper, i.e., to what extent the decline of marriage is accounted for by changes in the wage structure, I perform the following decomposition. The columns of tables 3.6 and 3.7 are labeled after the names of the experiments and provide an extensive overview of the decomposition.

- a In the first experiment, I simulate a counterfactual equilibrium where all parameters are fixed to their estimated 1970s levels except for the wage distribution, which is assigned its 2000s shape. More precisely, both the wage levels w_i and w_j and the transition matrices π_m and π_f are allowed to take on their 2000s values: in the next section, I will further decompose this step to understand how each of these components matters.
- b In the second experiment, I simulate a counterfactual equilibrium where all parameters are fixed to their estimated 1970s levels except for the shares of college graduates, which are assigned their 2000s values.
- c In the third experiment, I let both the wage distribution and the shares of college graduates correspond to what we observe in the 2000s. Both the production function of the public good, the meeting function and the cost of divorce are still fixed to their 1970s levels.
- d In the last experiment, I let the meeting function Λ and the cost of divorce κ take on the values indicated by their 2000s point-estimates in table 3.2. Only the parameters of the production function of the public good, γ_1 , γ_2 and γ_3 , are still fixed to their 1970s levels: these parameters can account for the residual change.

Different elements of the wage distribution may impact on the marriage market outcome differently. Hence, I run four additional simulations to explore the economic mechanism

that establishes the relationship between changes in the structure of labor market earnings and changes in marital patterns. Experiments are once again labeled by letters, consistent with the notation used in tables 3.8 and 3.9.

- e In experiment (e), I simulate a counterfactual equilibrium where all parameters are fixed to their estimated 1970s levels except for men’s wage levels, which are set to their 2000s values.
- f In experiment (f), I simulate a counterfactual equilibrium where all parameters are fixed to their estimated 1970s levels except for women’s wage levels, which are set to their 2000s values. Changes in men’s and women’s wage levels are plotted in figure 3.3 and 3.4.
- g In experiment (g), I simulate a counterfactual equilibrium where all parameters are fixed to their estimated 1970s levels except for the transition matrix π_m , which is allowed to take on its 2000s values. The elements of π_m determine the odds of changing wage quantile for men. Recall from section 3.5.1 that π_m is generated from an $AR(1)$ process and that the degree of wage mobility is monotonically decreasing in the wage rank correlation between an agent’s wage w in period t and w' in $t + 1$. Hence, looking at plot 3.5, it is easy to see that wage mobility among men has unequivocally increased between the 1970s and the 2000s.
- h In experiment (h), I simulate a counterfactual equilibrium where all parameters are fixed to their estimated 1970s levels except for the transition matrix π_f , which is allowed to take on its 2000s values. The elements of π_f determine the odds of changing wage quantile for women. Figure 3.5 indicates that wage mobility has decreased for women.

3.6.5. Overall changes in the wage distribution. Between the 1970s and 2000s, the share of married agents aged between 20 and 60 has fallen by 15.95 percentage points (i.e., a 21.4% decrease with respect to the initial level). Comparing the marriage market outcomes in the 1970s (column 1) with experiment (a), I conclude that changes in the wage structure can account for about 33.3% of the decline between the two periods (see table 3.6). Changes in the wage structure induce a larger decline in marriage among low-educated men (5.97pp) than among high-educated men (2.81pp); the decline is more evenly distributed among women: 5.48pp for those with a college degree and 4.23pp for those without. Marriage rates are lower while divorce rates are higher: the share of married agents grows less rapidly over the life-cycle and it reaches a lower peak with respect to the benchmark economy (see plot 3.11). The wage correlation among those who get married increases and the fraction of married women working full-time also grows.

Experiments (e) and (f) shed further light on the underlying mechanism: they show how changes in wage levels, for men and women respectively, have affected the marriage market

outcome while holding wage mobility fixed. Results are reported in tables 3.8 and 3.9. In experiment (e), only male wages change: male population groups that experience the strongest wage decline - the low-educated, the young and the low earners - are also those who experience the strongest decline in the odds of being married. In experiment (f), only female wages change: contrarily to experiment (e), female population groups that experience the strongest wage rise - the high-educated and the top earners - are those who experience the strongest decline in the odds of being married. In both experiments (e) and (f), the overall number of married agents decreases.

The findings from these experiments suggest that, when the gender wage gap shrinks, the economic gains from marriage tend to fade. As they experience wage rises, women can afford being more selective and value noneconomic gains more than in the past. Since the changes in wages have not been uniform across the population, some groups are particularly concerned by the evolution of the wage distribution. Consider the case of men without a college degree: while their wages decline, the gains from household specialization start to disappear. On the other hand, high-wage men are still able to build two kinds of family: as they retain their comparative advantage within the couple over low-wage women, they can form “traditional” households and enjoy the gains from specialization; alternatively, they can form couples of two earners with high-wage women, as the spouses’ pooled labor income is sufficiently large to take full advantage of the economies of scale and make the match convenient.

3.6.6. Changes in wage mobility. The last two columns of tables 3.8 and 3.9, (g) and (h), consider changes in wage mobility while holding wage levels fixed. This kind of experiment is, to the best of my knowledge, the first of its kind in the literature. My findings show that both changes in men’s wage mobility and women’s wage mobility have only a very small impact on the marriage market outcome. Men’s wage mobility have increased (see figure 3.5): column (g) in table 3.8 suggests that the effect of this increase has a positive - although very limited - impact on marriage. The explanation is that increased wage mobility makes marriage more relevant due to its insurance motive.

Women’s wage mobility is considered in experiment (h): this decrease has nearly no impact on the marriage market outcome. This can be explained as follows: at the 1970s equilibrium, changes in women’s wages almost never cause couples to break up, both because the wife can compensate a wage cut by increasing her domestic time and because the ratio of the wife’s earnings to the total household labor income is low.

3.6.7. Changes in schooling. Column (b) in table 3.6 suggests that the increase in the share of college graduates from the 1970s to the 2000s, which is particularly strong for women, has contributed to both a decline in marriage and changes in assortativeness. In experiment (b), the share of couples where both spouses are college graduates doubles,

and high-educated women start “marrying down”; in spite of this, their odds of being married decrease with respect to the 1970s benchmark. In experiment (c), changes in wages are considered jointly with changes in schooling. The interaction between the two changes leads to a further decrease in the share of married adults: these two factors can jointly account for about 49% of the decrease between the 1970s and the 2000s. With respect to experiment (a), changes in schooling partly offset the decrease in marriage for high-educated men.

Changes in the educational composition of the population only affect the match surplus through competition on the marriage market. Since the increase in the share of college graduates is weaker for men, high-educated women must be less selective. As a result, it becomes more difficult for them to find a valuable match. On the contrary, low-educated women are fewer, and are ready to match with both high-educated and low-educated men: their odds of being in a marriage relationship increase.

3.6.8. Changes in meetings and household production. Changes in the structure of meetings also contribute to the evolution of marriage patterns. Changes in the parameters of the search function are reported in table 3.2 and suggest that meetings are less assortative with respect to education in the 2000s. Experiment (d) shows that this results in a slight decline in the number of marriages (table 3.6) and in a decrease in the share of homogamous marriages (table 3.7). If agents are less likely to meet potential partners with a similar schooling level, they also tend to refuse more matches: this can also be seen by the higher number of singles among the young. However, since search is costly in terms of time, they become more inclined to accept marriages with people that are less alike.

As discussed in section 3.6.2, the cost of divorce is found to be increasing between the two periods, although it is high in both periods. This implies that, in all other experiments, most of the adjustments in the share of married across the population go through changes in marriage rates rather than in divorce rates. Since commitment is strong, individuals are extremely careful about selecting their partner and the search takes longer.

Changes in the parameters of the domestic production function explain the residual changes in marriage patterns, i.e., the difference between column (d) and column (2) in tables 3.6 and 3.7. In particular, they account for a residual 49% decline in the share of married adults. Three elements need to be considered: first, economies of scale have become stronger for couples where the wife holds a college degree, with a positive impact on the economic gains from marriage (see table 3.1). The opposite has happened for couples where the wife does not hold a college degree. Second, domestic production of public good has become less labor intensive, with a negative impact on the gains from specialization. For this reason, wives have stronger incentives to work and economies of

scale play a more important role than in the past: this favors high-wage women. Third, taste for age and educational homogamy has also changed. The vector γ_3 of structural parameters estimated in the current framework aims to capture additional motives for marriage not explicitly considered. Changes in these parameters are particularly relevant in explaining the changing age profile of marriage: the young have become comparatively less likely to be in a marital relationship.

3.6.9. Welfare analysis. In table 3.10, I report changes in the *ex-ante* welfare of individuals aged 20 for each gender group, before and after conditioning on education (see details on welfare measurement in section 3.2.13). Experiment (a) shows that changes in the wage distribution induce an average 9.5% welfare loss for men with respect to the 1970s benchmark economy. This welfare loss can be further decomposed: declining real wages at the middle and bottom of the distribution have a negative direct impact on human wealth and induce an average welfare loss of 8.2%; on top of that, worsened marriage prospects result in a 1.3% welfare loss due to reduced expected gains from marriage enjoyed over the life-cycle. As for women, experiment (a) shows that improved labor market conditions result in an average 1.3% welfare gain; however, this is offset by a 1.8% welfare loss due to worsened marriage prospects.

The bottom half of table 3.10 allows me to discuss the distributional effects of changes in the wage distribution across different educational groups. In experiment (a), welfare losses are far larger for men without a college degree (11.5%) than for college graduates (2.7%), although for both groups the losses are amplified through the marriage market. Female college graduates enjoy a sufficiently large increase in human wealth to offset the welfare loss due to lower expected gains from marriage and experience an overall welfare gain (4.3%). On the contrary, women without a college degree experience only minor improvements in labor market conditions: the gains associated with these improvements are offset by worsened marriage prospect and result in an overall welfare loss (1.4%). Finally, figure 3.12 documents changes in the distribution of *ex-ante* welfare for individuals with different initial wages.

In the last column of table 3.10, I report the welfare levels for the 2000s. On average, men have experienced welfare losses (8.6%): a considerable part of the losses is due to worsened marriage prospects (2.4%). Women have experienced net welfare gains (2.7%) and have partly caught up with men due to improved labor market conditions and increased schooling. However, lower expected marriage surplus has a negative impact on welfare (3.3%) and partly offsets the welfare increase realized outside of the marriage market (6%). As seen in experiment (a), the increase in human wealth for low-educated women is not sufficiently high to compensate the welfare losses due to worsened marriage prospects. On the contrary, female college graduates experience only a limited decrease

in expected marriage surplus. Figure 3.13 provides a graphical representation of this situation.

Between the 1970s and the 2000s, not only wage inequality has increased, but the marriage market has shifted from an equilibrium where the low and the high educated enjoy, on average, the same level of gains from marriage to an equilibrium where the high educated enjoy substantially more gains. The quantitative results presented in this section suggest that the marriage market amplifies economic inequality. One missing step is left for future work: in the current analysis, I suggest that the marriage market amplifies welfare differences across education and age groups, but the measure of welfare that I employ does not allow me to compare the finding with commonly used measures of economic inequality. A more satisfying measure requires to compute the monetary value of marriage surplus enjoyed by individuals on top of their human wealth. As anticipated in section 3.2.13, [Chiappori and Meghir \(2014\)](#) provide guidelines to compute measures of this kind in the case of static household models: their methodology can be extended to dynamic model in order to overcome these limitations.

3.7. Conclusion

In this paper, I build and estimate an equilibrium model of the marriage markets with search frictions, endogenous divorce, wage mobility and aging. This structural approach allows me to provide a quantitative assessment of the role of changes in the wage structure in explaining changes in the marriage market outcome. The empirical analysis is composed of the following steps: I first estimate the unobserved parameters of the model - the domestic production function, the meeting function and the cost of divorce - for both the 1970s and the 2000s. The estimated model is able to approximate the cross-sectional marital patterns - who is married and with whom - and the divorce and marriage rates observed in the data. I then proceed with a series of experiments where I analyze the role of changes in one primitive parameter of the model holding all other factors constant. I find that changes in the wage structure can explain about 33.3% of the decline in marriage that occurred between the 1970s and the 2000s, and that they have a stronger impact on the low educated. I also find that changes in positional inequality play a much more important role than changes in wage mobility. In particular, changes in men's wage inequality and the shrinking gender wage gap are the most important driving forces behind the decline of marriage. Finally, I show that in the 2000s, on top of the increased wage inequality, individuals with a college degree enjoy, on average, substantially more gains than those without: this gap in gains from marriage was instead absent in the 1970s. This result suggests that, as for the 2000s, the marriage market amplifies economic inequality.

The paper presents several innovative aspects, both in the modeling part and in the empirical analysis. It extends the search-and-matching framework of GJR by introducing

aging and wage mobility, and represents a first attempt to provide an empirically tractable framework to study marriage along the life-cycle and across cohorts.²⁹ The setup is potentially suitable for the analysis of a great variety of topics: the determinants of the age of first marriage, gender asymmetries in matching with respect to age and in remarriage trends, and the relationship between marriage and health. In the empirical analysis, I complement existing findings on marriage and economic inequality by providing a rich set of results, some of which are new to the literature. The analysis is the first to consider the joint impact of changes in wage inequality, wage mobility and the life-cycle dynamics of labor market earnings on the marriage market outcome. I show that changes in wage inequality between and within gender groups play a bigger role than changes in wage mobility. In section 3.6.8, I also briefly discuss the implications of changes in the way people meet each other on the marriage market outcome: while more work is needed in this direction, these first findings show that changes in the degree of segregation across educational groups matter in explaining the decline of marriage and the changing sorting patterns.

The paper represents a starting point for further research in this direction. In particular, the analysis abstracts away from human capital investment. Introducing choices such as schooling or dynamic labor supply decisions linked with human capital accumulation is key to understand how agents adjust to changes in labor market conditions. Previous works have stressed the importance of studying the interplay between human capital investment and competitive matching on the marriage market (e.g. [Chiappori, Salanié, and Weiss, 2017](#)); other works have extensively discussed the schooling and life-cycle career choices of women outside of an equilibrium framework (e.g. [Sheran, 2007](#); [Bronson, 2014](#)). Although this will add an additional layer of complexity, particularly in the estimation phase, the theoretical setup outlined in this paper can bridge these two literatures and provide new insights in this direction.

3.A. Technical Appendix

3.A.1. Linearizing the Pareto frontier. The TU property implies that there exist a cardinal representation of the spouses' preferences so that the Pareto frontier can be characterized as a straight line with slope -1 (see [Chiappori and Gugl, 2014](#)). This property is also used in the proof of lemma 1. Given the ordinal representation of spouses' utilities given by equation (3.2.1) and the demand for public good (3.2.5), it is possible to recover $\Gamma(l_f; i, j)$, i.e., the constant characterizing the linearized Pareto frontier associated

²⁹To the best of my knowledge, only [Shephard \(2018\)](#) is, at this date, working on this kind of models.

with l_f .

$$\begin{aligned}
\Gamma(l_f; i, j) &= \exp(\log(w_i - t_m) - \log Q^*(l_f; i, j)) + \exp(\log(l_f w_j - t_f) - \log Q^*(l_f; i, j)) = \\
&= (w_i + l_f w_j - t_m - t_f) Q^*(l_f; i, j) = \\
&= \frac{\gamma_1(i, j)^{\gamma_1(i, j)}}{(1 + \gamma_1(i, j))^{1 + \gamma_1(i, j)}} (w_i + l_f w_j)^{1 + \gamma_1(i, j)} \exp(\gamma_2(l_f; j) + \gamma_3(i, j) + \eta_{l_f}).
\end{aligned} \tag{3.A.1}$$

Note from the expression above how η_{l_f} tilts the Pareto frontier. Moreover, $\Gamma(l_f; i, j)$ is closely related to what is defined in section as the “per-period marital surplus” $h(l_f; i, j)$ corresponds to $\log(\Gamma(l_f; i, j))$.

3.A.2. Private consumption and sharing rule. The amount of private consumption is jointly determined by the amount of surplus produced by a match and the way it is shared between the spouses. Consider the wife’s Bellman equation (3.2.16): the splitting rule (3.2.10) tells us that her share of surplus must be exactly one half of the total. This restriction implied by the splitting rule allows me to back out the wife’s amount of private consumption. Interestingly, it is possible to write the ratio of the wife’s to the private husband’s consumption as follows:

$$\frac{w_j - t_f^*(l_f; i, j)}{w_i - t_m^*(l_f; i, j)} = \frac{\exp(V_f^0(j) - \beta\psi_f(j) \sum_{j'} V_f^0(j') \pi_f(j, j'))^2}{\exp(V_m^0(i) - \beta\psi_m(i) \sum_{i'} V_m^0(i') \pi_m(i, i'))^2}. \tag{3.A.2}$$

3.A.3. Steady-state equilibrium as a fixed point. The steady-state search equilibrium can be thought of as a fixed-point of an operator $n \rightarrow T_{ext}n$, with $n = (n_m, n_f)$. First, it is necessary to discretize the sets of types $|\mathcal{I}|$ and $|\mathcal{J}|$, as wage rates are continuous variables in the data. Hence, n is a vector of length $|\mathcal{I}| + |\mathcal{J}|$.

- (1) Start the iteration with $k = 0$ and a guess for n^k .
- (2) For given n^k , solve a fixed-point problem T_{int} given by equations (3.2.17), (3.2.25) and (3.2.26), in order to find $V^0 = (V_m^0, V_f^0)$ so that $V^0 = T_{int}V^0$.
- (3) Update α using (3.2.24).
- (4) Substitute the matrix α into the law of motion (3.2.32) and solve forwards for m .
- (5) Use the accounting equations (3.2.27) and (3.2.28) to compute n^{k+1} .
- (6) If $\Delta(n^k, n^{k+1}) < \epsilon$, keep n^{k+1} , otherwise set $n^k = \delta n^{k+1} + (1 - \delta)n^k$ and restart from step 2.

As anticipated in section 3.2.12, while there is no theoretical result ensuring existence and uniqueness of the equilibrium, iteration of the fixed-point operator leads to convergence to a vector n^* . Many simulations have brought to me to conclude that convergence to n^* is obtained regardless of the initial points chosen to start the algorithm and for a very broad choice of the numerous primitive parameters. The only caveat is that, at the last

step of the algorithm described above, I update the distribution n by taking a convex combination of the last two obtained vectors in the sequence: experience suggests that setting $0 < \delta < 1$ and in particular sufficiently close to 0 (I set it equal to 0.2) allows the algorithm to converge for almost any choice of the primitive parameters.

More theoretical guidance and a proof of existence and, possibly, uniqueness of the equilibrium would help understand the property of the fixed-point operator and could possibly help to design faster solution methods. This is left for future research.

3.A.4. Plackett’s copula and transition matrix. In order to obtain the transition matrices π_m and π_f , I characterize the $AR(1)$ wage process through a copula that links the wage rank of an individual across two consecutive periods. Consider the case of a man i : the joint CDF of his current wage rank $r_i(h_i, a_i)$ and his future wage rank $r_{i'}(h_{i'}, a_{i'})$ is given by the Plackett’s copula:

$$C_m(u, v | h_i, a_i) = \frac{1 + \theta(h_i, a_i)(u + v) - [1 + \theta(h_i, a_i)(u + v)^2 - 4\theta(h_i, a_i)(\theta(h_i, a_i) + 1)uv]^{1/2}}{2\theta(h_i, a_i)} \quad (3.A.3)$$

where the parameter $\theta(h_i, a_i)$ is such that the higher $\theta(h_i, a_i)$, the lower the mobility. In particular, [Nelsen \(2007, Chapter 5\)](#) shows that θ is a monotonically increasing function of the Spearman’s rank correlation coefficient. Dropping the arguments (h_i, a_i) for the sake of clarity, the two are related as follows,

$$\rho = \frac{2\theta + \theta^2 - 2(1 + \theta) \log(\theta + 1)}{\theta^2}. \quad (3.A.4)$$

3.A.5. Standardization of empirical frequencies. In order to produce a joint distribution of spouses characteristics that is consistent with the stationarity assumption and with the observed matching behavior, I apply the following transformation to the raw empirical frequencies of married and single agents by type. This appendix closely follows [Greenwood, Guner, Kocharkov, and Santos \(2016\)](#), which in turns draws from [Mosteller \(1968\)](#) and relies on the solution algorithm of [Sinkhorn and Knopp \(1967\)](#) outlined below.

Call $(m_{raw}, n_{m,raw}, n_{f,raw})$ the empirical frequencies as measured straight from the data: these are associated with the marginals $p_{m,raw}$ and $p_{f,raw}$ through accounting constraints (3.2.27) and (3.2.28). However, for my empirical analysis, I choose to work with marginals \hat{p}_m and \hat{p}_f , whose estimation is detailed in section 3.5.3. Starting from $(m_{raw}, n_{m,raw}, n_{f,raw})$, I compute empirical frequencies $(\hat{m}, \hat{n}_m, \hat{n}_f)$ - the “standardized contingency table” - as follows.

- (1) Start the iteration with $k = 0$ and set $m^k = m_{raw}$, $n_m^k = n_{m,raw}$, $n_f^{k+1} = n_{f,raw}$.
- (2) Compute the marginal distribution p_m^k using m^k and n_m^k and accounting restriction (3.2.27).

- (3) For each man's type i , rescale each element of the contingency table as follows: $m^{k+1}(i, j) = m^k(i, j)(\hat{p}_m(i)/p_m^k(i))$ and $n^{k+1}(i) = n^k(i)(\hat{p}_m(i)/p_m^k(i))$, where \hat{p}_m is the men's marginal distribution that has been imposed.
- (4) Compute the marginal distribution p_f^{k+1} using m^{k+1} and n^{k+1} and accounting restriction (3.2.28).
- (5) For each woman's type j , rescale each element of the contingency table as follows: $m^{k+2}(i, j) = m^{k+1}(i, j)(\hat{p}_f(j)/p_f^k(j))$ and $n^{k+2}(j) = n^k(i)(\hat{p}_f(j)/p_f^k(j))$, where \hat{p}_f is the women's marginal distribution that has been imposed.
- (6) Compute p_m^{k+2} using m^{k+2} and n_m^k : if p_m^{k+2} and \hat{p}_m are close, stop; otherwise repeat from step 2 until convergence.

3.B. Tables

Table 3.1. Estimates of γ_1 by spouses' education

Parameter (λ_1)	1971-1981 (a)	2001-2011 (b)	Change (b)-(a)
Husband L, wife L	1.286 (0.098)	0.949 (0.080)	-0.336
Husband L, wife H	1.302 (0.143)	1.552 (0.095)	0.250
Husband H, wife L	1.546 (0.129)	1.157 (0.103)	-0.390
Husband H, wife H	1.002 (0.130)	1.559 (0.101)	0.557

Notes: the table contains estimates of γ_1 , the elasticity of public good Q with respect to the joint expenditure $t_m + t_f$, for different types of couples; L (Low) indicates that the agent does not hold a college degree, while H (High) indicates that he/she does hold a college degree. The first column contains results for the 1970s sample, the second for the 2000s sample, and the third the difference between the two. Standard errors are in parentheses.

Table 3.2. Estimates of the search parameters

Parameter	1971-1981 (a)	2001-2011 (b)	Change (b)-(a)
Educational homophily (λ_1)	0.645 (0.003)	0.361 (0.006)	-0.284
Age distance squared (λ_2)	-0.044 (0.001)	-0.005 (0.002)	0.039
Age term of degree 3 $a_m a_f^2$ (λ_3)	-0.004 (0.000)	-0.010 (0.000)	-0.006
Age term of degree 3 $a_m^2 a_f$ (λ_4)	-0.001 (0.000)	0.007 (0.000)	0.008
CES parameter (χ)	100.000	100.000	0.000
Cost of divorce (κ)	4.221 (0.001)	4.623 (0.001)	0.402

Notes: the table contains estimates of the parameters of the meeting function Λ and the cost of divorce κ . The elasticity of substitution of inputs ($n_{m,+}(i), n_{f,+}(j)$) in the meeting function χ is set to 100 after that early findings suggested that the meeting function is well approximated by a Leontief function. The empirical specification of the meeting function is detailed in section 3.5.6.

Table 3.3. Summary of parameters estimated or calibrated outside of the model

Parameter	1971-1981	2001-2011	Notes
β	0.9604	0.9604	Calibrated (Voena, 2015; Chiappori, Salanié, and Weiss, 2017)
π	0.7	0.7	Calibrated (see section 3.5.5)
Share of college graduates (men)	20.84%	30.54%	Estimated (CPS data, see section 3.5.3)
Share of college graduates (women)	13.03%	31.31%	Estimated (CPS data, see section 3.5.3)
Wage levels (men and women)	See figure 3.3		Estimated (CPS data, see section 3.5.1)
Wage mobility (men and women)	See figure 3.3		Estimated (PSID data, see section 3.5.1)

Notes: the table summarizes the parameters that are calibrated or estimated outside of the model. Wage levels and wage mobility are described in separate figures.

Table 3.4. Fit of the model (1)

Moment	1971-1981		2001-2011		Change	
	(a)		(b)		(b) - (a)	
	Sim.	Data	Sim.	Data	Sim.	Data
<i>% of married</i>						
Men	74.59	69.74	58.64	52.14	-15.95	-17.60
Women	74.59	67.69	58.64	52.45	-15.95	-15.24
<i>% of married by education</i>						
Men (L)	76.03	68.78	57.26	47.18	-18.77	-21.60
Women (L)	75.39	67.84	56.83	48.93	-18.56	-18.91
Men (H)	69.09	73.79	61.77	65.03	-7.32	-8.75
Women (H)	69.23	66.65	62.60	60.99	-6.63	-5.67
<i>% of married by wage quintile</i>						
Men (1st quintile)	61.63	49.31	43.35	29.50	-18.27	-19.81
Women (1st quintile)	73.11	60.50	49.36	38.98	-23.75	-21.52
Men (3rd quintile)	75.01	72.44	53.43	48.39	-21.58	-24.05
Women (3rd quintile)	74.29	73.02	52.26	52.39	-22.03	-20.64
Men (5th quintile)	86.80	87.48	74.57	76.60	-12.23	-10.88
Women (5th quintile)	79.43	65.97	68.67	62.43	-10.76	-3.54
<i>% of married by age</i>						
Men (20-30)	41.44	45.59	20.65	20.36	-20.79	-25.23
Women (20-30)	59.19	57.10	30.17	30.02	-29.02	-27.09
Men (30-40)	80.34	78.69	60.43	56.83	-19.91	-21.86
Women (30-40)	85.51	76.70	67.78	61.00	-17.72	-15.70
Men (40-50)	87.17	82.47	72.87	63.27	-14.30	-19.20
Women (40-50)	88.47	77.09	74.98	62.50	-13.49	-14.59
Men (50-60)	88.73	82.96	78.61	67.33	-10.12	-15.62
Women (50-60)	74.69	67.50	67.35	57.62	-7.35	-9.88

Notes: the table contains moments from the empirical distribution of agents' marital status conditional on their characteristics for both the 1970s and the 2000s. These are compared to the moments obtained by simulating the estimated model. In the last two columns, it is possible to assess the changes observed in the data and those implied by the two simulations. Labels: (L) means non-college graduate; (H) means college graduate; (xx-yy) means from age xx to age yy; (PT) means part-time; (FT) means full-time.

Table 3.5. Fit of the model (2)

Moment	1971-1981		2001-2011		Change	
	(a)		(b)		(b) - (a)	
	Sim.	Data	Sim.	Data	Sim.	Data
<i>% of couples by education</i>						
L husband, L wife	79.55	76.12	61.15	54.92	-18.40	-21.20
L husband, H wife	1.15	3.58	6.68	10.42	5.53	6.84
H husband, L wife	8.36	11.39	5.43	11.12	-2.93	-0.27
H husband, H wife	10.95	8.91	26.75	23.53	15.80	14.62
<i>Correlation between spouses' traits</i>						
Wage	20.05	25.87	26.19	33.57	6.14	7.70
Wage (rank)	13.96	27.12	23.88	34.87	9.92	7.75
Age	69.51	75.41	65.69	67.66	-3.83	-7.75
<i>% of couples by wife's l_f</i>						
Not working (L)	41.71	49.25	26.34	29.28	-15.37	-19.97
Working PT (L)	20.73	16.37	22.06	17.82	1.33	1.44
Working FT (L)	37.56	34.38	51.60	52.91	14.04	18.53
Not working (H)	34.74	35.10	22.31	20.77	-12.43	-14.33
Working PT (H)	18.24	17.53	17.69	17.25	-0.55	-0.28
Working FT (H)	47.02	47.37	60.00	61.98	12.98	14.61
<i>Marriage rates</i>						
Men	18.78	26.81	11.30	10.95	-7.48	-15.86
Women	20.48	20.94	11.78	10.39	-8.70	-10.55
Men (L)	20.24	27.67	10.92	10.35	-9.31	-17.32
Women (L)	21.30	21.09	11.25	9.37	-10.05	-11.71
Men (H)	14.39	25.22	12.22	12.09	-2.17	-13.13
Women (H)	16.06	20.56	13.09	12.08	-2.96	-8.48
<i>Divorce rates</i>						
Overall	1.82	5.09	2.25	4.91	0.43	-0.18
Men (L)	1.88	5.45	2.60	6.24	0.72	0.79
Women (L)	1.81	5.21	2.45	6.01	0.64	0.80
Men (H)	1.56	4.20	1.50	3.21	-0.07	-0.98
Women (H)	1.91	4.64	1.84	3.50	-0.07	-1.14

Notes: the table contains moments from the joint empirical frequency of spouses' characteristics, married women's empirical probabilities of choosing a certain level l_f of labor supply and the distribution of empirical probabilities of getting married (for singles) or divorced (for married) for both the 1970s and the 2000s. These are compared to the moments obtained by simulating the estimated model. In the last two columns, it is possible to assess the changes observed in the data and those implied by the two simulations. Labels: (L) means non-college graduate; (H) means college graduate; (xx-yy) means from age xx to age yy; (PT) means part-time; (FT) means full-time.

Table 3.6. Decomposition of the changes in marriage market outcome (1)

	1970s	Experiments				2000s
	(1)	Wages (a)	Schooling (b)	Wages Schooling (c)	Wages Schooling Meetings (d)	(2)
<i>% of married</i>						
Overall	74.59	69.27	72.16	66.73	66.51	58.64
<i>% of married by education</i>						
Men (L)	76.03	70.06	73.63	66.45	66.30	57.26
Women (L)	75.39	69.91	76.01	69.58	70.18	56.83
Men (H)	69.09	66.28	68.82	67.37	67.01	61.77
Women (H)	69.23	65.00	63.72	60.48	58.48	62.60
<i>% of married by wage quintile</i>						
Men (1st quintile)	61.63	56.75	56.03	50.72	50.59	43.35
Women (1st quintile)	73.11	70.14	69.42	65.24	65.02	49.36
Men (3rd quintile)	75.01	70.82	69.83	63.42	63.03	53.43
Women (3rd quintile)	74.29	72.59	73.44	63.75	64.01	52.26
Men (5th quintile)	86.80	82.25	86.64	81.26	81.38	74.57
Women (5th quintile)	79.43	70.80	74.24	68.70	68.39	68.67
<i>% of married by age</i>						
Men (20-30)	41.44	36.84	37.77	33.14	30.67	20.65
Women (20-30)	59.19	55.25	54.70	50.52	49.09	30.17
Men (30-40)	80.34	74.29	78.30	72.07	70.27	60.43
Women (30-40)	85.51	79.90	83.61	77.62	76.42	67.78
Men (40-50)	87.17	81.50	85.24	79.49	80.08	72.87
Women (40-50)	88.47	82.81	86.17	80.45	81.46	74.98
Men (50-60)	88.73	83.70	86.33	81.21	83.24	78.61
Women (50-60)	74.69	68.46	73.35	67.23	68.01	67.35

Notes: the table contains moments from different simulated economies: column (1) aims to match the 1970s equilibrium; the labels of the column in the middle part of the table (“Experiments”) refers to the names of the experiments described in section 3.6.4; column (2) aims to match the 2000s equilibrium. Each experiment corresponds to a simulated equilibrium where all parameters are fixed to their 1970s levels but those named in the header, which are set to their 2000s levels. The table helps understand which factors contributed the most to changes in the marriage market outcome. Finally, note that the residual difference between column (c) and (2) is only explained by changes in the production function of the public good Q (i.e., due to changes in γ_1 , γ_2 and γ_3).

Table 3.7. Decomposition of the changes in marriage market outcome (2)

	1970s	Experiments				2000s
	(1)	Wages (a)	Schooling (b)	Wages Schooling (c)	Wages Schooling Meetings (d)	(2)
<i>% of couples by education</i>						
L husband, L wife	79.55	79.30	65.71	65.35	64.28	61.15
L husband, H wife	1.15	0.76	5.16	3.82	4.95	6.68
H husband, L wife	8.36	8.47	6.64	6.28	8.19	5.43
H husband, H wife	10.95	11.47	22.49	24.56	22.58	26.75
<i>Correlation between spouses' traits</i>						
Wage	20.05	21.07	23.58	25.40	23.26	26.19
Wage (rank)	13.96	16.16	18.47	22.25	20.29	23.88
Age	69.51	67.55	70.08	67.88	66.06	65.69
<i>% of couples by wife's l_f</i>						
Not working (L)	41.71	37.52	41.76	37.61	37.69	26.34
Working PT (L)	20.73	21.05	20.70	21.02	21.04	22.06
Working FT (L)	37.56	41.43	37.54	41.37	41.27	51.60
Not working (H)	34.74	31.02	33.96	30.14	29.64	22.31
Working PT (H)	18.24	18.28	18.18	18.22	18.19	17.69
Working FT (H)	47.02	50.70	47.86	51.64	52.17	60.00
<i>Marriage rates</i>						
Men	18.78	16.22	17.51	15.10	13.87	11.30
Women	20.48	17.45	18.96	16.13	14.90	11.78
Men (L)	20.24	17.04	18.35	15.34	13.96	10.92
Women (L)	21.30	18.04	21.77	18.65	17.65	11.25
Men (H)	14.39	13.44	15.76	14.54	13.69	12.22
Women (H)	16.06	14.05	14.26	11.85	10.54	13.09
<i>Divorce rates</i>						
Overall	1.82	2.44	2.12	2.71	1.90	2.25
Men (L)	1.88	2.55	2.22	2.94	2.04	2.60
Women (L)	1.81	2.46	2.07	2.78	1.86	2.45
Men (H)	1.56	1.99	1.89	2.19	1.60	1.50
Women (H)	1.91	2.32	2.28	2.53	2.00	1.84

Notes: the table contains moments from different simulated economies: column (1) aims to match the 1970s equilibrium; the labels of the column in the middle part of the table (“Experiments”) refers to the names of the experiments described in section 3.6.4; column (2) aims to match the 2000s equilibrium. Each experiment corresponds to a simulated equilibrium where all parameters are fixed to their 1970s levels but those named in the header, which are set to their 2000s levels. The table helps understand which factors contributed the most to changes in the marriage market outcome. Finally, note that the residual difference between column (c) and (2) is only explained by changes in the production function of the public good Q (i.e., due to changes in γ_1 , γ_2 and γ_3).

Table 3.8. Changes in wage mobility and wage levels (1)

	1970s	Experiments			
	(1)	Men's wages (e)	Women's wages (f)	Men's mobility (g)	Women's mobility (h)
<i>% of married</i>					
Overall	74.59	70.11	73.70	74.75	74.60
<i>% of married by education</i>					
Men (L)	76.03	70.82	75.25	76.17	76.05
Women (L)	75.39	70.47	74.76	75.55	75.40
Men (H)	69.09	67.42	67.83	69.35	69.10
Women (H)	69.23	67.73	66.64	69.40	69.24
<i>% of married by wage quintile</i>					
Men (1st quintile)	61.63	56.71	60.70	62.29	61.64
Women (1st quintile)	73.11	68.63	73.97	73.26	73.34
Men (3rd quintile)	75.01	71.85	74.17	75.07	75.03
Women (3rd quintile)	74.29	69.54	77.26	74.45	74.32
Men (5th quintile)	86.80	83.88	86.00	86.47	86.81
Women (5th quintile)	79.43	75.21	75.51	79.61	79.19
<i>% of married by age</i>					
Men (20-30)	41.44	36.82	41.23	41.52	41.46
Women (20-30)	59.19	55.21	58.95	59.28	59.21
Men (30-40)	80.34	74.83	79.70	80.45	80.36
Women (30-40)	85.51	80.48	84.85	85.63	85.52
Men (40-50)	87.17	82.65	86.10	87.32	87.19
Women (40-50)	88.47	84.10	87.27	88.65	88.49
Men (50-60)	88.73	85.21	87.28	89.02	88.73
Women (50-60)	74.69	70.01	73.21	74.97	74.70

Notes: the table contains moments from different simulated economies: column (1) aims to match the 1970s equilibrium; the labels of the column in the middle part of the table (“Experiments”) refers to the names of the experiments described in section ???. Each experiment corresponds to a simulated equilibrium where all parameters are fixed to their 1970s levels but those named in the header, which are set to their 2000s levels. The table helps understand the impact of changes in each element of the wage distribution on the changing marriage market outcome.

Table 3.9. Changes in wage mobility and wage levels (2)

	1970s	Experiments			
	(1)	Men's wages (e)	Women's wages (f)	Men's mobility (g)	Women's mobility (h)
<i>% of couples by education</i>					
L husband, L wife	79.55	79.23	79.63	79.53	79.55
L husband, H wife	1.15	0.73	1.19	1.14	1.15
H husband, L wife	8.36	8.18	8.59	8.37	8.36
H husband, H wife	10.95	11.86	10.59	10.96	10.95
<i>Correlation between spouses' traits</i>					
Wage	20.05	21.10	19.16	20.03	20.12
Wage (rank)	13.96	15.37	13.60	13.92	14.04
Age	69.51	67.45	69.54	69.50	69.52
<i>% of couples by wife's l_f</i>					
Not working (L)	41.71	38.34	40.87	41.67	41.72
Working PT (L)	20.73	21.02	20.79	20.74	20.73
Working FT (L)	37.56	40.64	38.34	37.60	37.55
Not working (H)	34.74	33.68	32.13	34.70	34.75
Working PT (H)	18.24	18.27	18.26	18.24	18.24
Working FT (H)	47.02	48.06	49.61	47.06	47.01
<i>Marriage rates</i>					
Men	18.78	16.50	18.37	18.88	18.79
Women	20.48	17.81	19.96	20.61	20.49
Men (L)	20.24	17.31	19.83	20.34	20.24
Women (L)	21.30	18.23	20.95	21.44	21.31
Men (H)	14.39	13.75	14.01	14.49	14.40
Women (H)	16.06	15.23	14.86	16.12	16.06
<i>Divorce rates</i>					
Overall	1.82	2.25	1.99	1.80	1.82
Men (L)	1.88	2.37	2.04	1.87	1.88
Women (L)	1.81	2.29	1.96	1.79	1.81
Men (H)	1.56	1.79	1.77	1.52	1.56
Women (H)	1.91	1.97	2.24	1.88	1.91

Notes: the table contains moments from different simulated economies: column (1) aims to match the 1970s equilibrium; the labels of the column in the middle part of the table (“Experiments”) refers to the names of the experiments described in section 3.6.4. Each experiment corresponds to a simulated equilibrium where all parameters are fixed to their 1970s levels but those named in the header, which are set to their 2000s levels. The table helps understand the impact of changes in each element of the wage distribution on the changing marriage market outcome.

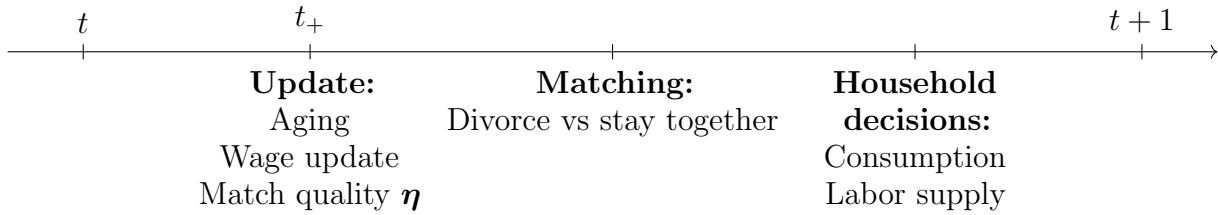
Table 3.10. Changes in the distribution of intertemporal welfare at the beginning of adulthood

	1970s	Experiments				2000s
	(1)	Wages (a)	Schooling (b)	Wages Schooling (c)	Wages Schooling Meetings (d)	(2)
<i>By gender</i>						
Average welfare (men)	100.00	90.50	101.64	93.15	92.62	91.40
Average welfare without marriage (men)	86.32	78.13	87.56	80.06	80.06	80.06
Average expected marriage gains (men)	13.68	12.37	14.07	13.08	12.56	11.33
Average welfare (women)	81.50	81.09	84.93	85.60	85.18	83.69
Average welfare without marriage (women)	66.30	67.36	69.23	71.20	71.20	71.20
Average expected marriage gains (women)	15.20	13.73	15.69	14.39	13.98	12.49
<i>By gender and education</i>						
Average welfare (men, L)	97.38	86.17	97.02	85.92	85.32	84.00
Average welfare without marriage (men, L)	83.65	73.97	83.65	73.97	73.97	73.97
Average expected marriage gains (men,L)	13.74	12.20	13.37	11.95	11.35	10.03
Average welfare (men, H)	109.93	106.95	112.15	109.59	109.21	108.23
Average welfare without marriage (men, H)	96.47	93.92	96.47	93.92	93.92	93.92
Average expected marriage gains (men,H)	13.47	13.02	15.68	15.67	15.29	14.31
Average welfare (women, L)	79.04	77.94	80.28	79.23	78.89	75.81
Average welfare without marriage (women, L)	64.21	64.62	64.21	64.62	64.62	64.62
Average expected marriage gains (women, L)	14.83	13.32	16.07	14.61	14.27	11.18
Average welfare (women, H)	97.97	102.13	95.12	99.56	98.99	100.98
Average welfare without marriage (women, H)	80.25	85.63	80.25	85.63	85.63	85.63
Average expected marriage gains (women, H)	17.72	16.50	14.87	13.93	13.36	15.35

Notes: the table reports the evolution of the intertemporal welfare of individuals entering the marriage market at age \underline{a} between the two periods and provides a decomposition of the change. The *ex-ante* welfare is decomposed into two terms according to equation (3.2.35): the first accounts for the welfare generated by the agents' human wealth outside of the marriage market, the second for the expected marriage gains over the life-cycle (see section 3.2.13 for details). All measures are normalized so that the average welfare for men in the 1970s is equal to 100. A full description of the experiments from (a) to (d) can be found in section 3.6.4.

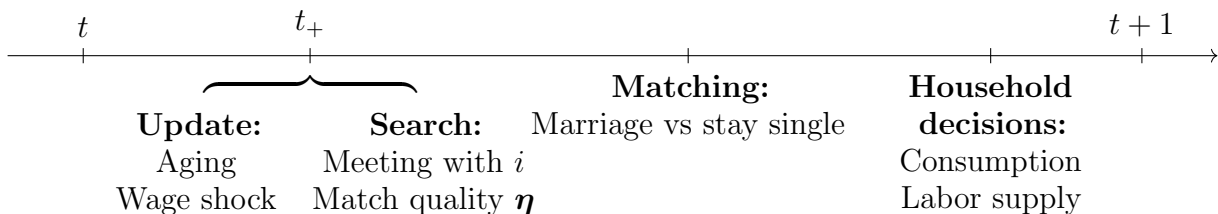
3.C. Figures

Figure 3.1. Timeline for a man i and a woman j married in t



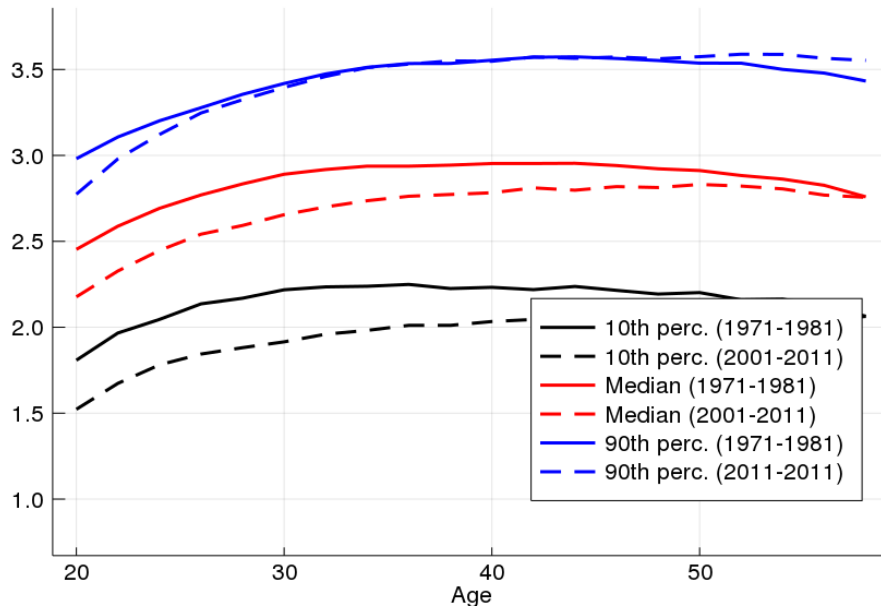
Notes: the figure reproduces the timing of decisions of agents that enter the period as married couples. Uncertainty is resolved at the beginning of the period (t_+): both spouses draw new wages and experience taste shocks η . Immediately after t_+ , couples have sufficient information to decide whether to stay together or divorce. Conditionally on their updated marital status, they make consumption and labor supply decisions.

Figure 3.2. Timeline for a woman j who is single in t

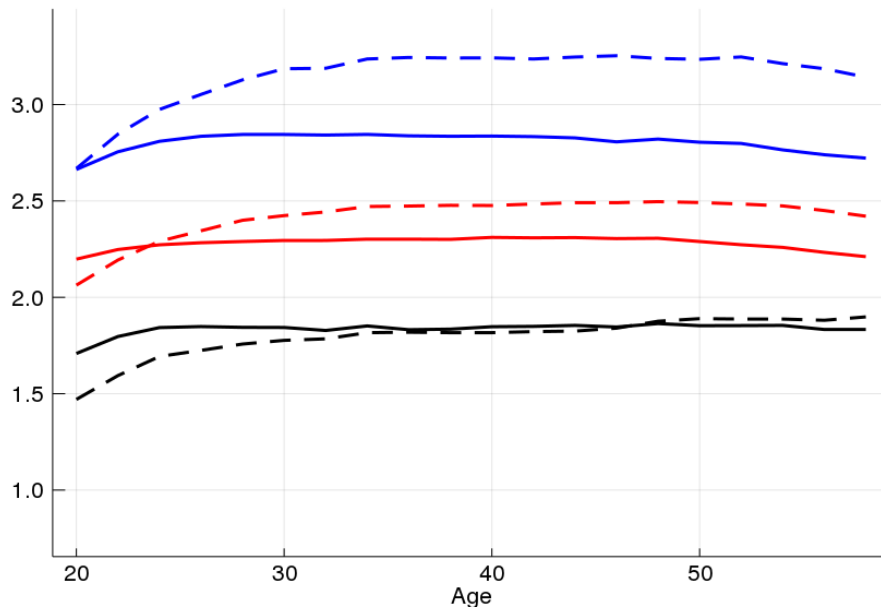


Notes: the figure reproduces the timing of single agents' decisions. Uncertainty is resolved at the beginning of the period (t_+): singles draw a new wage and look for a partner; upon a date, the pair draws a vector η which is indicative of match quality. Immediately after t_+ , single agents have sufficient information to decide whether to get married (if they have met someone). Conditionally on their updated marital status, they make consumption and labor supply decisions.

Figure 3.3. Wage levels by gender and age



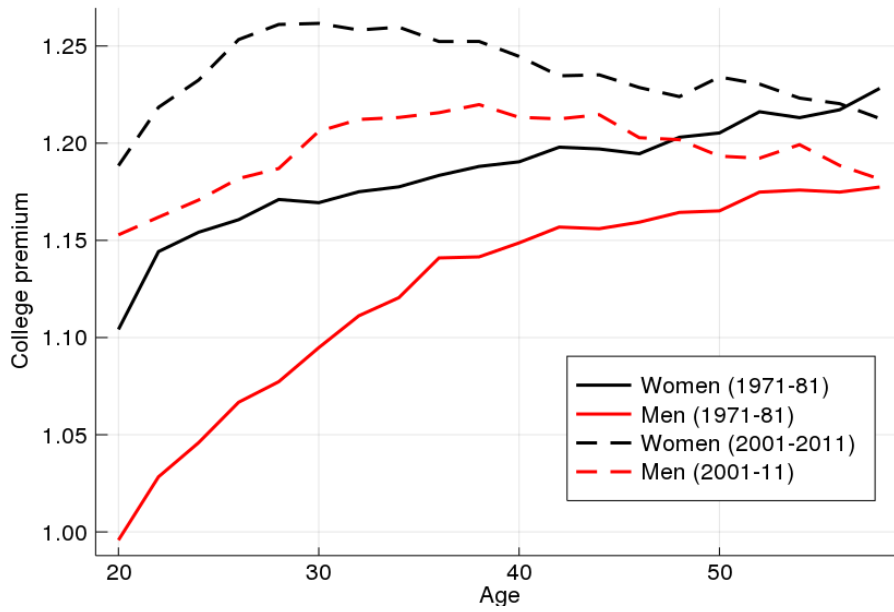
(a) Men



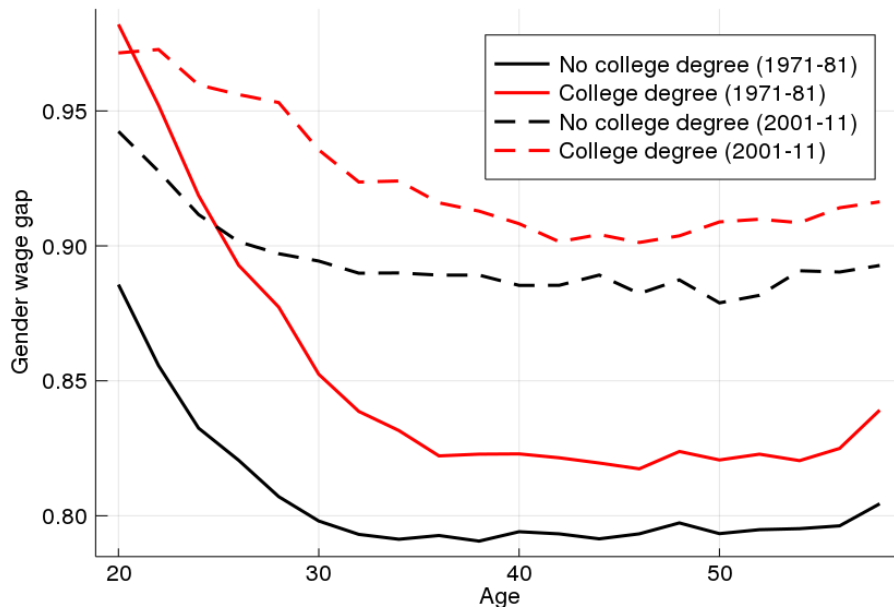
(b) Women

Notes: the plots display wage levels for the 10th (black), 50th (red) and 90th percentile (blue) for both men (on the left) and women (on the right). Solid lines refer to wage levels in the 1970s, while dash lines refer to the 2000s.

Figure 3.4. College premium and gender wage gap



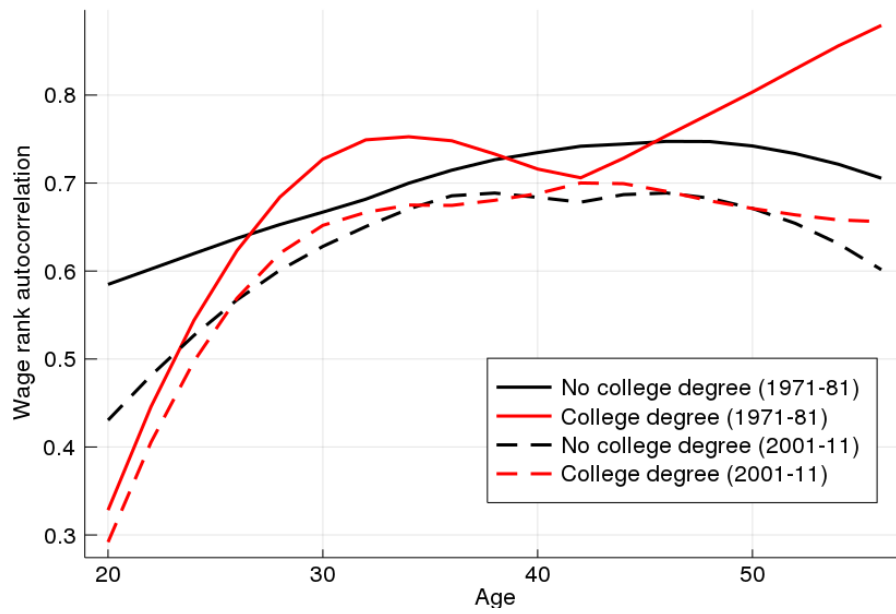
(a) College premium



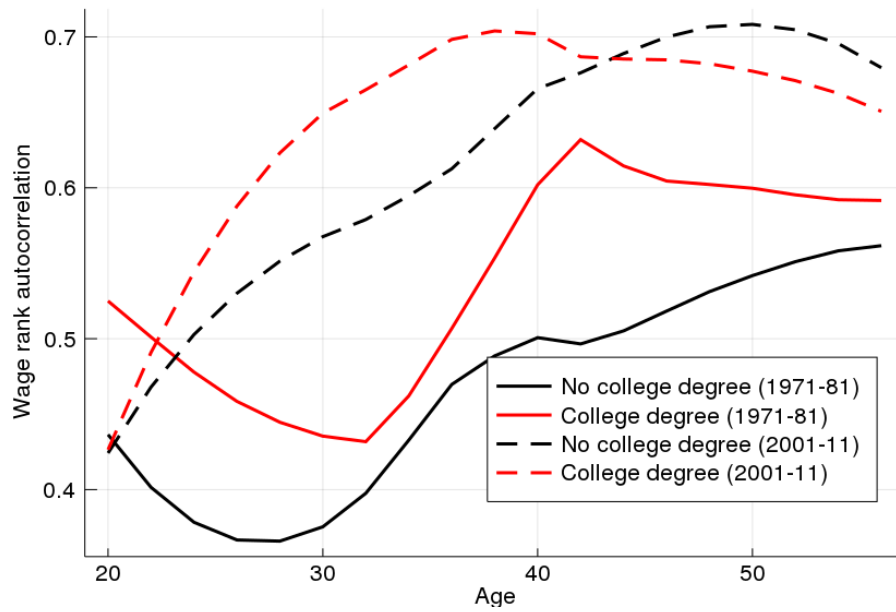
(b) Gender wage gap

Notes: the plots display the college premium (on the right) and the gender wage gap (on the left) measured as ratios between the median wages of each gender and educational group. Solid lines refer to ratios in the 1970s, while dash lines refer to the 2000s.

Figure 3.5. Changes in wage mobility by gender, age and education



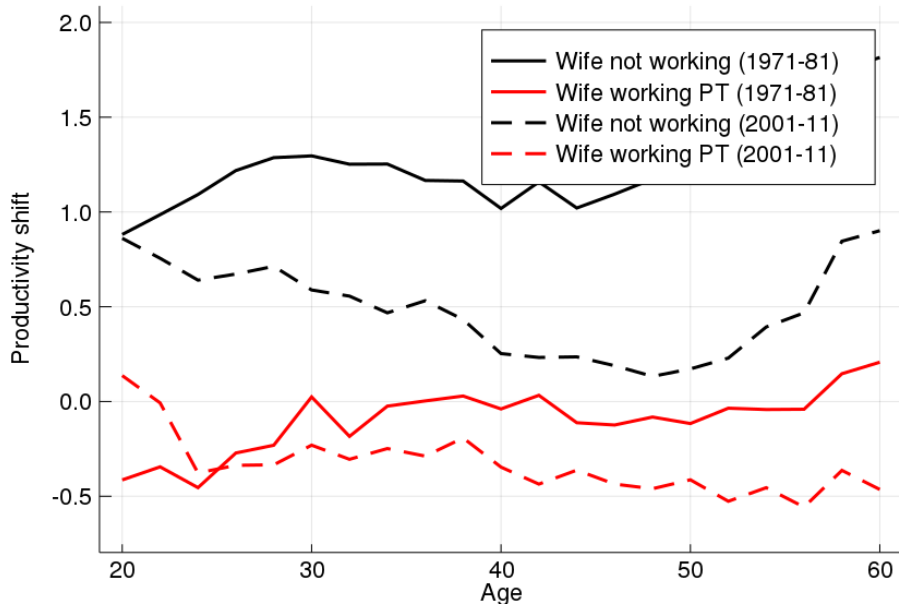
(a) Men



(b) Women

Notes: the plots display the rank correlation between the wage w in period t and the wage w' in period $t + 1$: the higher the rank correlation, the lower wage mobility. Solid lines refer to parameter estimates for the 1970s, while dash lines refer to the 2000s.

Figure 3.6. Estimates of γ_2 by age and education



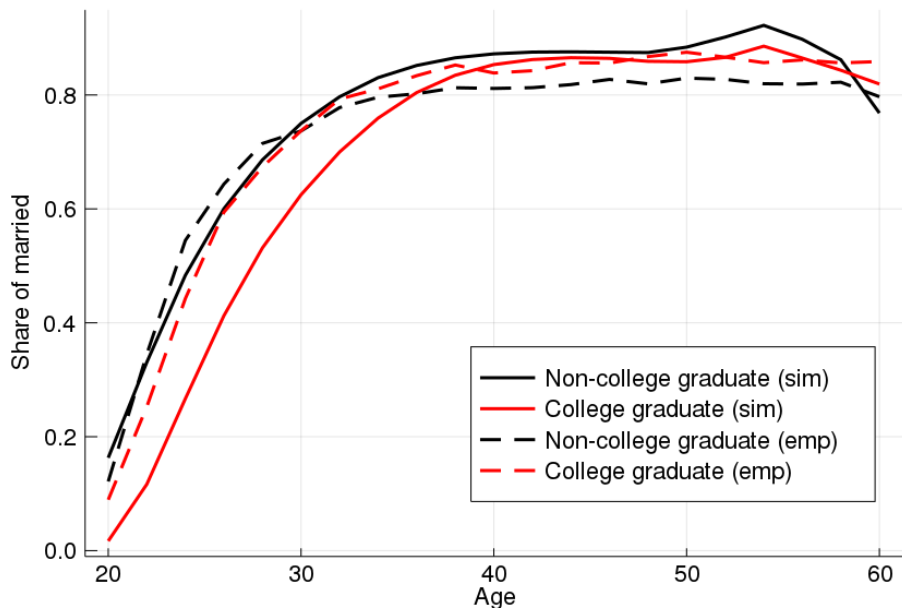
(a) Noncollege graduates



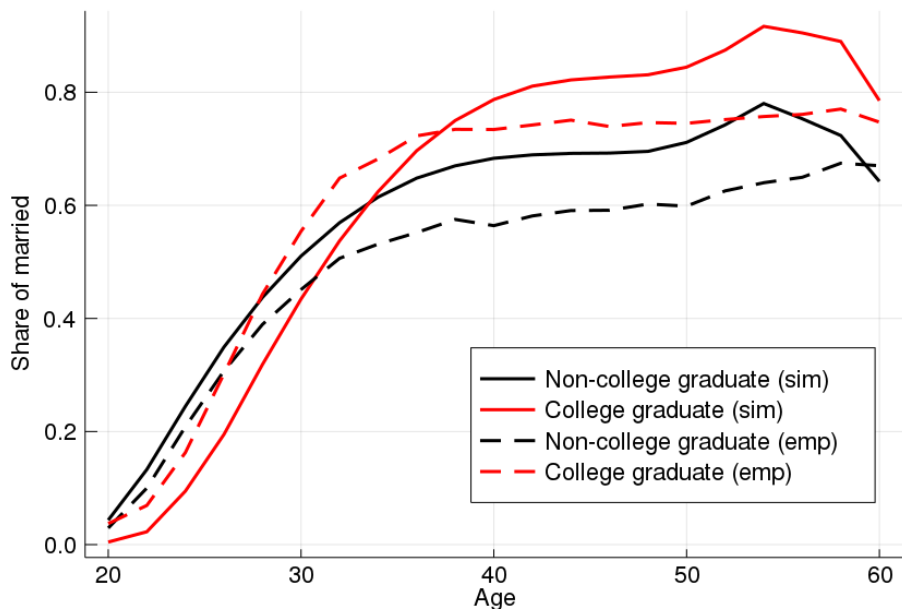
(b) College graduates

Notes: the plots display estimates of $\gamma_2(l_f)$, the productivity shifter associated with labor supply choice l_f , for women without a college degree (on the left) and with a college degree (on the right). Black indicates inactivity and red part-time; solid lines represent 1970s estimates, and dashed lines 2000s estimates. For each level l_f , $\gamma_2(l_f)$ is allowed to vary by age and education; $\gamma_2(l_f)$ is normalized to zero for married women working full-time. If $\gamma_2(l_f) > 0$, the household benefits from an increase in match surplus with respect to the benchmark (wife working full-time) if option l_f is chosen; if $\gamma_2(l_f) < 0$, the household faces a loss in match surplus.

Figure 3.7. Fit of the model: share of married men by age and education



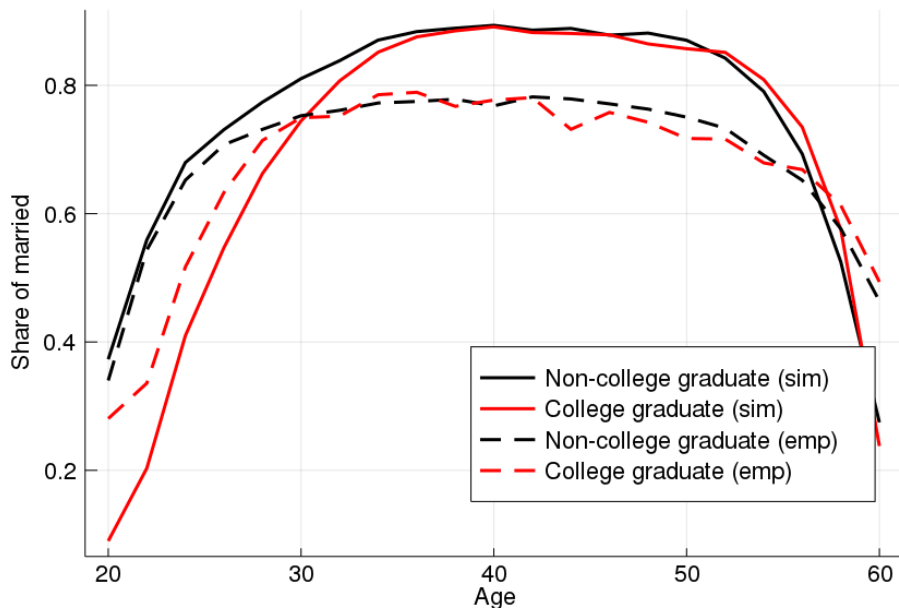
(a) 1970s



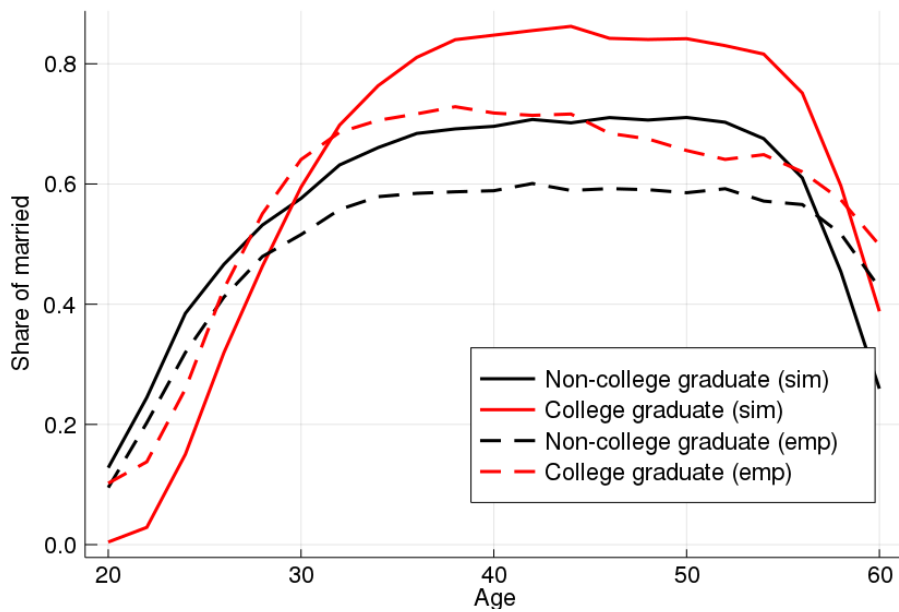
(b) 2000s

Notes: the plots track the life-cycle dynamics of the stocks of married men by educational level. For each plot, black represents non-college graduates, and red college graduates; solid lines correspond to simulated moments, dash lines to empirical moments.

Figure 3.8. Fit of the model: share of married women by age and education



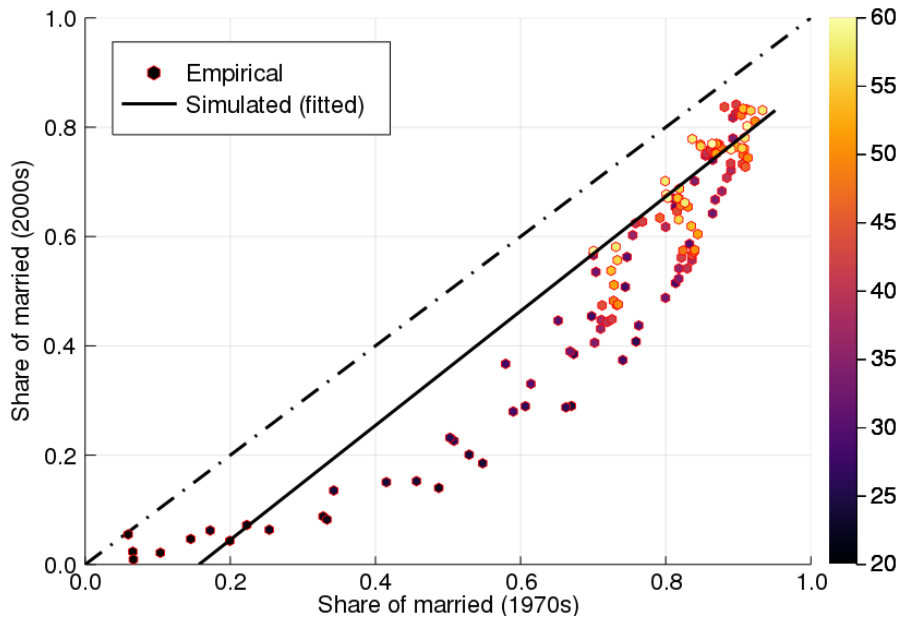
(a) 1970s



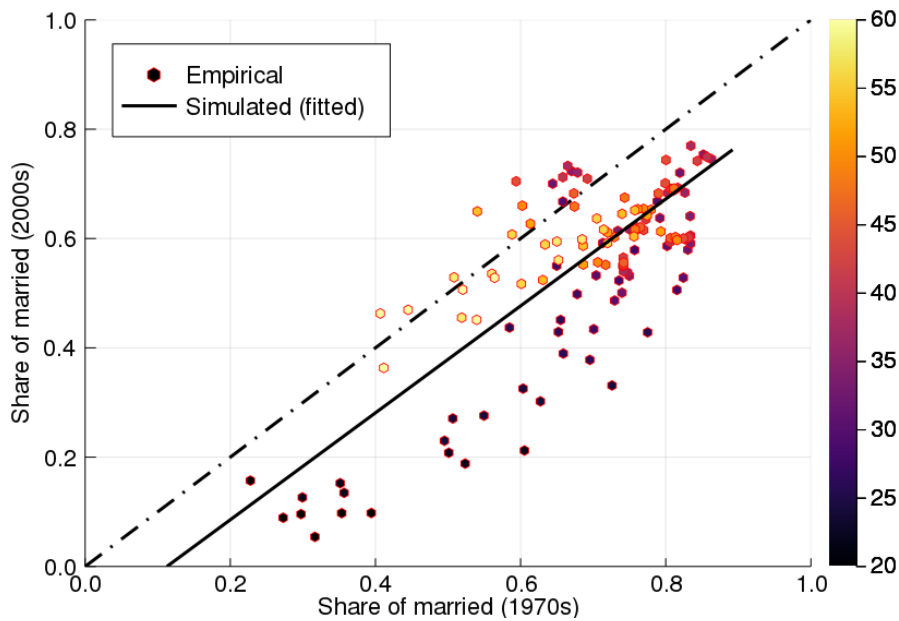
(b) 2000s

Notes: the plots track the life-cycle dynamics of the stocks of married men by educational level. For each plot, black represents non-college graduates, and red college graduates; solid lines correspond to simulated moments, dash lines to empirical moments.

Figure 3.9. Fit of the model: changes on the extensive margin by gender and age



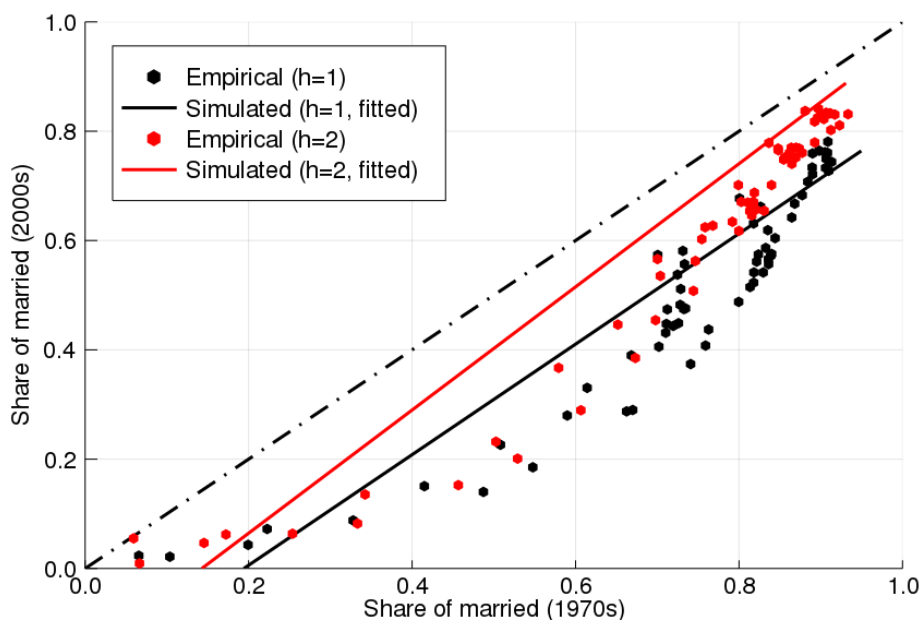
(a) Changes in the share of single men by age



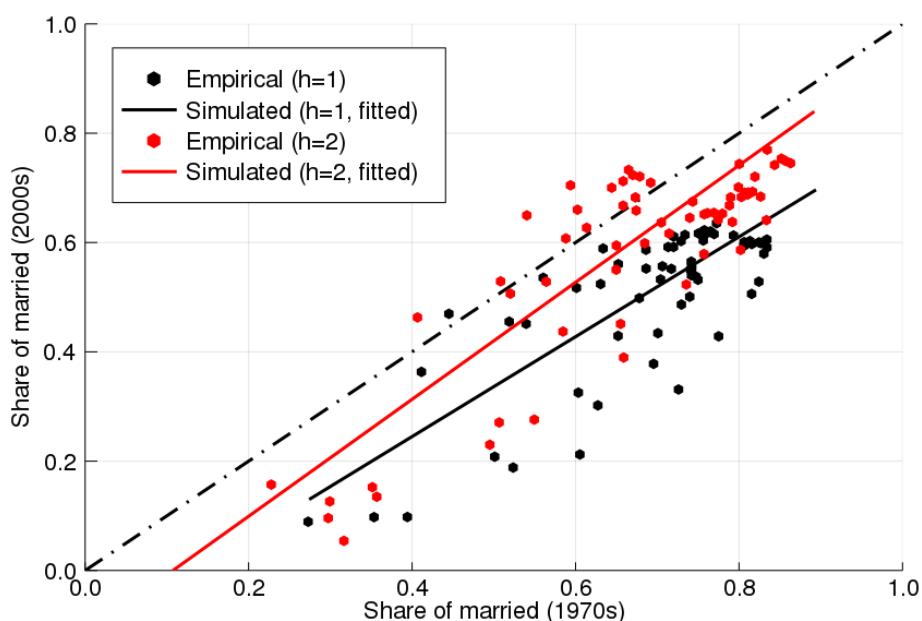
(b) Changes in the share of single women by age

Notes: the scatter plots reports the share of married agents in the 2000s vs the share of married agents in the 1970s as observed in the data for a given type i (i.e., by education, age and wage quantile). Markers are colored according to the age associated with the type (see colorbar on the right). Points below the 45-degree line correspond to a decline of marriage for a given i . The solid black line corresponds to the fitted values obtained by regressing the predicted shares of married agents in the 2000s on the predicted shares of married agents in the 1970s.

Figure 3.10. Fit of the model: changes on the extensive margin by gender and education



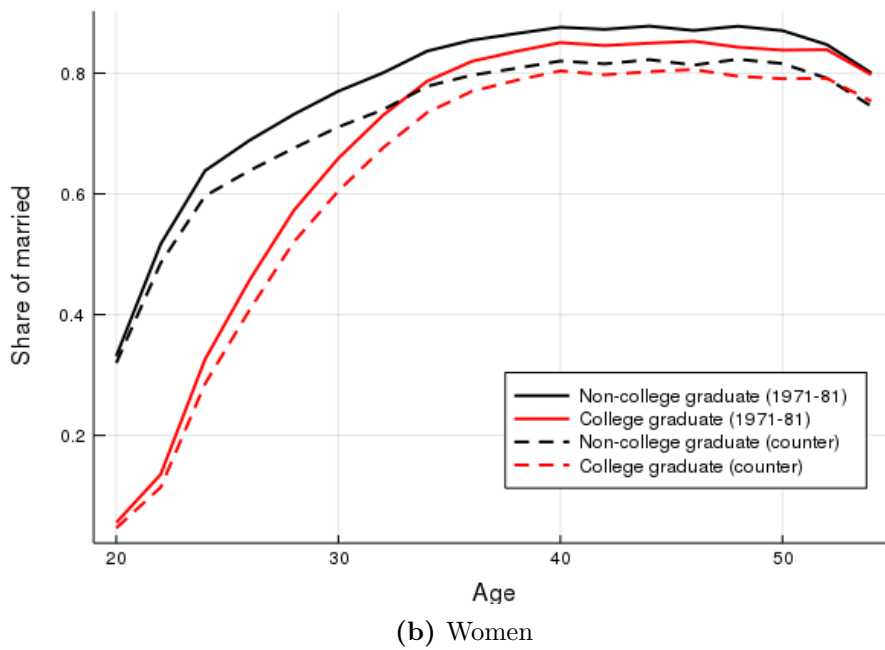
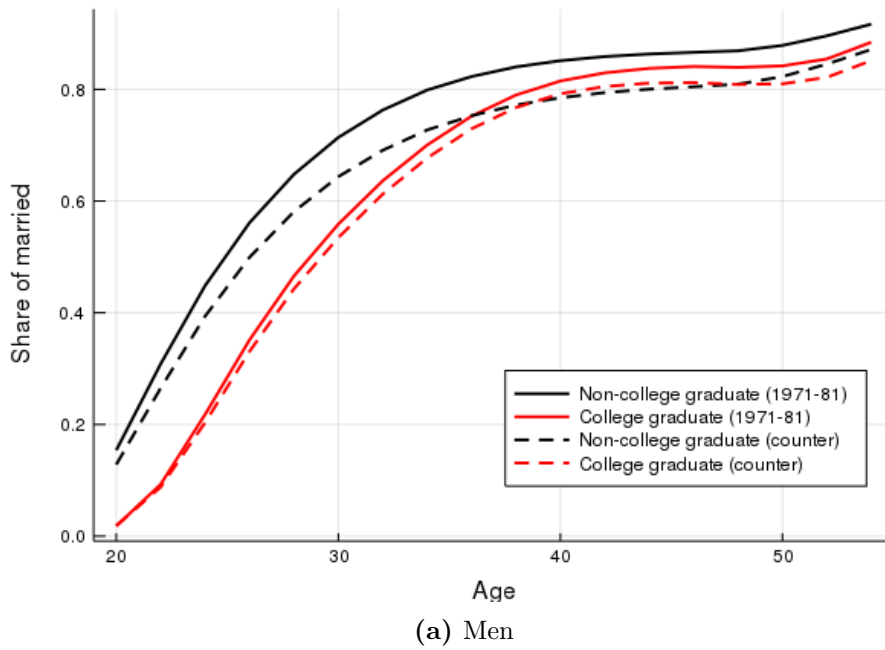
(a) Changes in the share of single men by education



(b) Changes in the share of single women by education

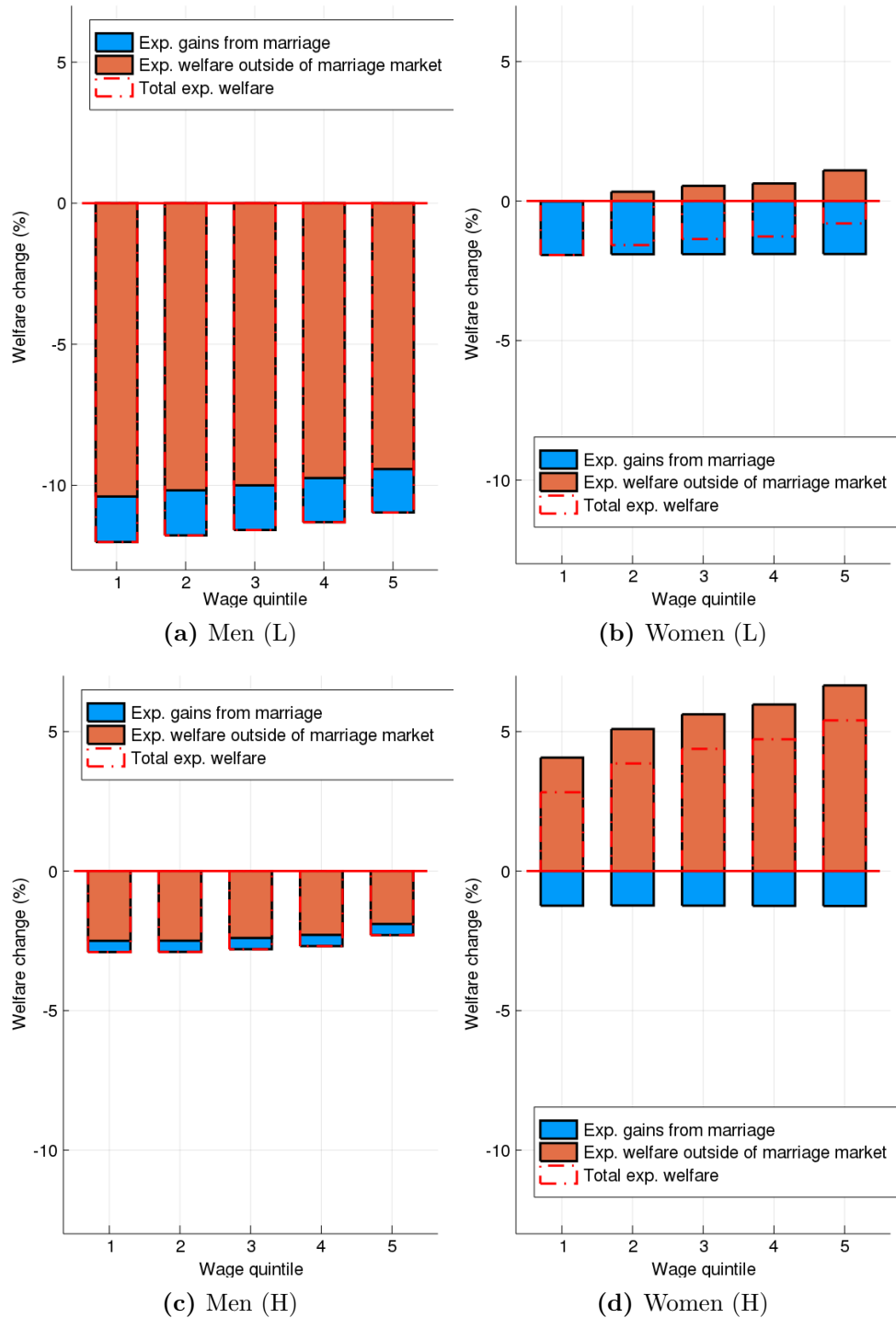
Notes: the scatter plots reports the share of married agents in the 2000s vs the share of married agents in the 1970s as observed in the data for a given type i after separating the population by gender and educational group ($h = 2$ indicates a college degree). Points below the 45-degree line correspond to a decline of marriage for a given i . The solid lines correspond to the fitted values obtained by regressing the predicted shares of married agents in the 2000s on the predicted shares of married agents in the 1970s for a given gender and educational group.

Figure 3.11. Changes in the share of married men by age and education: experiment (a)



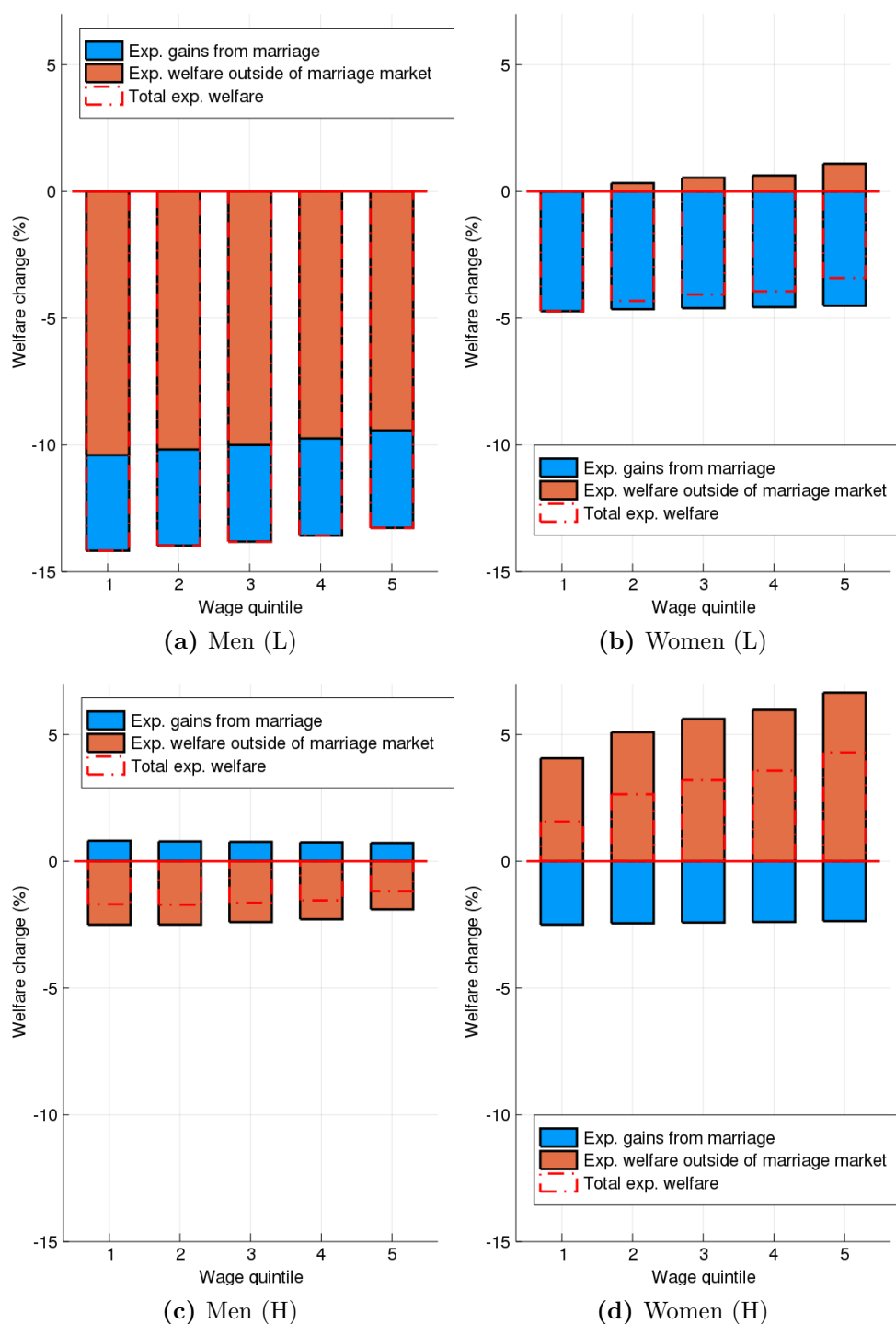
Notes: the plots track the life-cycle dynamics of the stocks of married men by educational level. For each plot, black represents non-college graduates, and red college graduates; solid lines correspond to simulated moments for the 1970s marriage market (column (1) in table 3.6), while dash lines correspond to simulated moments for a counterfactual equilibrium where all parameters are fixed to the 1970s but the wage distribution, which takes on its 2000s shape (column (a) in table 3.7).

Figure 3.12. Changes in intertemporal welfare of young individuals: experiment (a)



Notes: the plots report changes in ex-ante welfare for individuals entering the marriage market at age 20 in the context of experiment (a) (see section 3.2.13). Blue bars correspond to % change in expected gains from marriage over the life-cycle, while orange bars correspond to changes in the welfare generated outside of the marriage market. The dash red bars correspond to the net change obtained as the sum of the two other terms.

Figure 3.13. Changes in intertemporal welfare of young individuals: from the 1970s to the 2000s



Notes: the plots report changes in ex-ante welfare for individuals entering the marriage market at age 20 when considering the full transition from the 1970s to the 2000s (see section 3.2.13). Blue bars correspond to % change in expected gains from marriage over the life-cycle, while orange bars correspond to changes in the welfare generated outside of the marriage market. The dash red bars correspond to the net change obtained as the sum of the two other terms.

Bibliography

- ACEMOGLU, D., AND D. AUTOR (2011): “Skills, tasks and technologies: Implications for employment and earnings,” *Handbook of labor economics*, 4, 1043–1171.
- ADDA, J., C. DUSTMANN, AND K. STEVENS (2017): “The career costs of children,” *Journal of Political Economy*, 125(2), 293–337.
- AIYAGARI, S. R., J. GREENWOOD, AND N. GUNER (2000): “On the State of the Union,” *Journal of Political Economy*, 108(2), 213–244.
- BEAUCHAMP, A., G. SANZENBACHER, S. SEITZ, AND M. M. SKIRA (2018): “Single moms and deadbeat dads: The role of earnings, marriage market conditions, and preference heterogeneity,” *International Economic Review*, 59(1), 191–232.
- BECKER, G. S. (1973): “A Theory of Marriage: Part I,” *Journal of Political Economy*, 81(4), 813–846.
- BLUNDELL, R., M. C. DIAS, C. MEGHIR, AND J. M. SHAW (2016): “Female Labour Supply, Human Capital and Welfare Reform,” *Econometrica*, p. forthcoming.
- BONHOMME, S., AND J.-M. ROBIN (2009): “Assessing the equalizing force of mobility using short panels: France, 1990–2000,” *The Review of Economic Studies*, 76(1), 63–92.
- BRONSON, M. A. (2014): “Degrees are forever: Marriage, educational investment, and lifecycle labor decisions of men and women,” .
- BROWNING, M., AND P.-A. CHIAPPORI (1998): “Efficient intra-household allocations: A general characterization and empirical tests,” *Econometrica*, pp. 1241–1278.
- BRUZE, G., M. SVARER, AND Y. WEISS (2015): “The dynamics of marriage and divorce,” *Journal of Labor Economics*, 33(1), 123–170.
- CHADE, H., AND G. VENTURA (2002): “Taxes and marriage: a two-sided search analysis,” *International Economic Review*, 43(3), 955–985.
- CHIAPPORI, P.-A. (1988): “Rational Household Labor Supply,” *Econometrica*, 56(1), 63–90.
- (1992): “Collective Labor Supply and Welfare,” *Journal of Political Economy*, 100(3), 437–67.
- CHIAPPORI, P.-A., M. C. DIAS, AND C. MEGHIR (forthcoming): “The Marriage Market, Labor Supply and Education Choice,” *The Journal of Political Economy*.
- CHIAPPORI, P.-A., AND E. GUGL (2014): “Necessary and sufficient conditions for transferable utility,” Discussion paper, Mimeo, Columbia University.
- CHIAPPORI, P.-A., AND M. MAZZOCCO (2017): “Static and intertemporal household decisions,” *Journal of Economic Literature*, 55(3), 985–1045.
- CHIAPPORI, P.-A., AND C. MEGHIR (2014): “Intra-household welfare,” Discussion paper, National Bureau of Economic Research.
- CHIAPPORI, P.-A., B. SALANIÉ, AND Y. WEISS (2017): “Partner choice, investment in children, and the marital college premium,” *American Economic Review*, 107(8), 2109–67.

- CHOO, E. (2015): “Dynamic marriage matching: An empirical framework,” *Econometrica*, 83(4), 1373–1423.
- CHOO, E., AND A. SIOW (2006): “Who Marries Whom and Why,” *Journal of Political Economy*, 114(1), 175–201.
- (2007): “Lifecycle Marriage Matching: Theory and Evidence,” Discussion paper, Society for Economic Dynamics.
- CISCATO, E. (2018): “Marriage, Divorce and Wage Uncertainty along the Life-cycle,” Discussion Paper 2018-046, Human Capital and Economic Opportunity Working Group, <https://hceconomics.uchicago.edu/research/working-paper/marriage-divorce-and-wage-uncertainty-along-life-cycle>.
- DEMUYNCK, T., AND T. POTOMS (2018): “Weakening Transferable Utility: the Case of Non-intersecting Pareto Curves,” *ECARES Working Papers*, 2018.
- DUPUY, A., AND A. GALICHON (2014): “Personality traits and the marriage market,” *Journal of Political Economy*, 122(6), 1271–1319.
- ECKSTEIN, Z., AND O. LIFSHITZ (2011): “Dynamic female labor supply,” *Econometrica*, 79(6), 1675–1726.
- ECKSTEIN, Z., AND K. I. WOLPIN (1989): “Dynamic labour force participation of married women and endogenous work experience,” *The Review of Economic Studies*, 56(3), 375–390.
- FERNÁNDEZ, R., N. GUNER, AND J. KNOWLES (2005): “Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality,” *Quarterly Journal of Economics*, 120(1), 273–344.
- FERNÁNDEZ, R., AND R. ROGERSON (2001): “Sorting and Long-Run Inequality,” *Quarterly Journal of Economics*, 116(4), 1305–1341.
- FERNÁNDEZ, R., AND J. C. WONG (2017): “Free to leave? A welfare analysis of divorce regimes,” *American Economic Journal: Macroeconomics*, 9(3), 72–115.
- FLABBI, L., AND C. FLINN (2015): “Simultaneous Search in the Labor and Marriage Markets with Endogenous Schooling Decisions,” .
- GALICHON, A., AND B. SALANIÉ (2015): “Cupid’s invisible hand: Social surplus and identification in matching models,” .
- GEMICI, A. (2011): “Family migration and labor market outcomes,” *Manuscript, New York University*.
- GOUSSÉ, M. (2014): “Marriage Market and Intra-Household Allocation,” Ph.D. thesis, Sciences Po Paris.
- GOUSSÉ, M., N. JACQUEMET, AND J.-M. ROBIN (2017a): “Household labour supply and the marriage market in the UK, 1991-2008,” *Labour Economics*, 46, 131–149.
- (2017b): “Marriage, labor supply, and home production,” *Econometrica*, 85(6), 1873–1919.
- GREENWOOD, J., N. GUNER, AND J. A. KNOWLES (2003): “More on marriage, fertility, and the distribution of income,” *International Economic Review*, 44(3), 827–862.

- GREENWOOD, J., N. GUNER, G. KOCHARKOV, AND C. SANTOS (2014): “Marry Your Like: Assortative Mating and Income Inequality,” *The American Economic Review*, 104(5), 348.
- (2016): “Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation,” *American Economic Journal: Macroeconomics*, 8(1), 1–41.
- HECKMAN, J. (1974): “Shadow prices, market wages, and labor supply,” *Econometrica: journal of the econometric society*, pp. 679–694.
- KEANE, M. P., AND K. I. WOLPIN (2007): “Exploring the usefulness of a nonrandom holdout sample for model validation: Welfare effects on female behavior*,” *International Economic Review*, 48(4), 1351–1378.
- (2010): “The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women*,” *International Economic Review*, 51(3), 851–892.
- KOPCZUK, W., E. SAEZ, AND J. SONG (2010): “Earnings inequality and mobility in the United States: evidence from social security data since 1937,” *The Quarterly Journal of Economics*, 125(1), 91–128.
- LIGON, E., J. P. THOMAS, AND T. WORRALL (2002): “Informal insurance arrangements with limited commitment: Theory and evidence from village economies,” *The Review of Economic Studies*, 69(1), 209–244.
- LISE, J., AND J.-M. ROBIN (2017): “The macrodynamics of sorting between workers and firms,” *American Economic Review*, 107(4), 1104–35.
- LUNDBERG, S., AND R. A. POLLAK (2014): “Cohabitation and the uneven retreat from marriage in the United States, 1950–2010,” in *Human capital in history: The American record*, pp. 241–272. University of Chicago Press.
- MANEA, M. (2017): “Steady states in matching and bargaining,” *Journal of Economic Theory*, 167, 206–228.
- MAZZOCCO, M. (2007): “Household intertemporal behaviour: A collective characterization and a test of commitment,” *The Review of Economic Studies*, 74(3), 857–895.
- MAZZOCCO, M., C. RUIZ, AND S. YAMAGUCHI (2017): “Labor Supply, Wealth Dynamics, and Marriage Decisions,” *UCLA CCPR Population Working Papers*.
- MEGHIR, C., AND L. PISTAFERRI (2011): “Earnings, consumption and life cycle choices,” in *Handbook of labor economics*, vol. 4, pp. 773–854. Elsevier.
- MOSTELLER, F. (1968): “Association and estimation in contingency tables,” *Journal of the American Statistical Association*, 63(321), 1–28.
- NELSEN, R. B. (2007): *An introduction to copulas*. Springer Science & Business Media.
- POLLAK, R. A. (1990): “Two-sex demographic models,” *Journal of Political Economy*, 98(2), 399–420.
- REGALIA, F., AND J.-V. RIOS-RULL (2001): “What accounts for the increase in the number of single households?,” *University of Pennsylvania, mimeo*.

- REYNOSO, A. (2017): “The impact of divorce laws on the equilibrium in the marriage market,” .
- ROGERSON, R., R. SHIMER, AND R. WRIGHT (2005): “Search-theoretic models of the labor market: A survey,” *Journal of economic literature*, 43(4), 959–988.
- SHAPLEY, L. S., AND M. SHUBIK (1971): “The Assignment Game I: The Core,” *International Journal of Game Theory*, 1(1), 111–130.
- SHEPHARD, A. (2018): “Marriage market dynamics, gender, and the age gap,” .
- SHERAN, M. (2007): “The career and family choices of women: A dynamic analysis of labor force participation, schooling, marriage, and fertility decisions,” *Review of Economic Dynamics*, 10(3), 367–399.
- SHIMER, R., AND L. SMITH (2000): “Assortative Matching and Search,” *Econometrica*, 68(2), 343–369.
- SINKHORN, R., AND P. KNOPP (1967): “Concerning nonnegative matrices and doubly stochastic matrices,” *Pacific Journal of Mathematics*, 21(2), 343–348.
- TRAIN, K. E. (2009): “GEV,” in *Discrete choice methods with simulation*, chap. 4. Cambridge University Press.
- VAN DER KLAUW, W. (1996): “Female labour supply and marital status decisions: A life-cycle model,” *The Review of Economic Studies*, 63(2), 199–235.
- VOENA, A. (2015): “Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?,” *American Economic Review*, 105(8), 2295–2332.
- WIEBE, T. (2018): “Aging And The Gains From Marriage,” Ph.D. thesis, University of Pennsylvania.
- WONG, L. Y. (2003): “Structural Estimation of Marriage Models,” *Journal of Labor Economics*, 21(3), 699–727.