Single by Choice or Rejection? Evidence on Mating Preferences in China

Xuchao Gao*

November 12, 2024

Link to The Most Recent Version

Abstract

Some common mating preferences fail to rationalize the matching patterns observed in the Chinese marriage where women's single rates increase in education attainment but in contrast, more educated men are less likely to be single. I propose a utility model that allows testing for the intensity of traditional gender norms, specifically their preferences over a marriage with the wife being more educated. Based on the estimation with 2010 Census data from China, I find that that both men and women do not like the woman being more educated in the marriage, although men's distaste is more intense. The distastes from both genders shape the single rate patterns. Removing the aversion from either side would greatly reduce the single rates but implications on matching depend on which side abandons the aversion. Additionally, the shift in education distribution alone well explains the rising singles rates of the least education men but not for the most educated women.

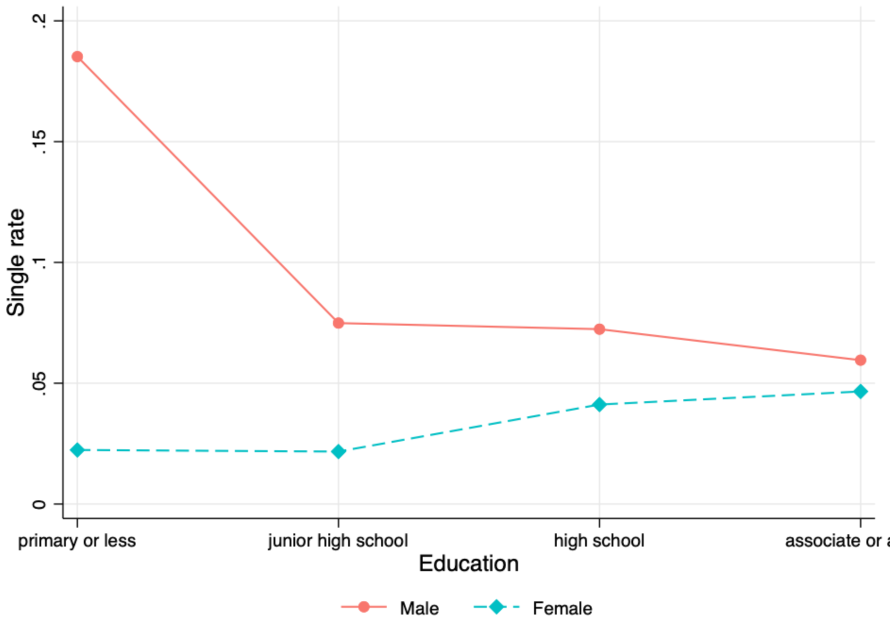
Keywords: marriage, matching, preferences, gender

*Email:xcgao@bu.edu. I am grateful to my main advisor Kevin Lang for his guidance and support. Without him this project would not have been possible. I also would like to thank my other advisors Daniele Paserman, Benjamin Marx, and other participants at the seminar in Boston University for helpful conversation and comments.

1 Introduction

Single rates have risen sharply among high-educated women and low-educated men in China. In the meanwhile, there is a striking divergence between men and women in how single rates change with education: for men, the likelihood of remaining single decreases, while for women, it increases. This pattern become clearer in recent cohorts. See the birth cohort of 1975 in Figure 1 as an example. For older cohorts, the gaps in the single rates for different education groups are much smaller, especially for women. This leads us to ponder why education seems to reward men in the marriage market but not helpful for women.

Figure 1: Single rates: males and females born in 1975



Note: This figure summarizes the single rates for females and males who were born in 1975, i.e. 35 years old at the time of survey. All individuals born in 1975, not in school at the time of survey, and were married, divorced, or widowed with a non-missing value for marital status are included. Those who married for the first time after 35 are treated as single in the calculation. Data source: Census 2010 1% sample.

In this paper, I take the education distributions as given and ask what mating preferences would explain the marriage pattern of people born in 1975. I then show that these same preferences combined with the education distributions in earlier periods can explain much of the large increases in the proportions of low-education men who are single.

To estimate the preferences, I employ a flexible utility model that allows for a mixture of assortative matching and homogamy. The non-transferable utility that an individual gets from marriage does not only depend on the spouse's level of education but also on the difference between their education levels in some circumstances. Specifically, I assume that the difference in education levels affects their utilities when the woman's education is higher.

The reluctance of both women and men—or of one side—to enter a marriage where the woman has higher education is consistent with the distinctive patterns observed in the single rates of the two genders. If men are resistant to marrying women with more education than themselves, more educated men will have more options in the marriage market and, therefore, exhibit lower single rates. On the other hand, if women are unwilling to marry men with lower education, the most educated women may find themselves in a thin marriage market. These implications align with the data. In conjunction with the more significant educational advancement among women in recent decades, this reluctance may help explain the rise in single rates for well-educated women, which mass media often refers to as "left-behind women" phenomenon.

This aversion is deeply rooted in traditional culture and has started to draw attention in research. The social norm that "men are breadwinners, women are homemakers" (Chen and Hu (2021)) has not kept up to women's empowerment in education over the past few decades. Chen and Hu (2021) indeed found that the couples are less satisfied-especially the husband-with their marriage if the wife earns more, based on Chinese household survey data from 2014. With U.S. data, Bertrand et al. (2015) also documented similar aversion to a marriage with the wife earning more, which they show affects the marriage formation, women's labor supply and undertaking of household chores as well as the couple's satisfaction of the marriage. This prejudice could also exist in non-economic aspects like education. This paper provides evidence on the existence of this aversion along education dimension in the mating preferences, estimates the intensities of the aversion for men and women, and using the estimated preferences, offers an explanation for the contrasting patterns between men and women in their single rate variations with education.

Based on the utility model and a simulated sample that retains the proportions for each education and gender group in the Census, I adopt the deferred acceptance algorithm to match men and women, in which men propose to women. Due to the complexity of the matching process, Newton Ralphson method or ther derivative-based methods do not apply. I use the Nelder-Mead algorithm to perform the optimization.

The estimation results indicate that both women and men generally prefer not to be in a relationship where the woman is more educated than the man, although the distaste among men is stronger than among women. This mutual distaste generates the relationship between education levels and singlehood for two genders in the Census data. Furthermore, reducing the aversion from either gender could lower the single rates among highly educated women and less-educated men; however, the impact on their marriage prospects depends on which side's aversion is addressed.

This paper contributes to the literature of marriage sorting by estimating a novel mating preference that rationalizes the distinctive patterns in single rates for men and women. Researchers have been interested in studying assortative mating along dimensions like income, education, or something more of inherited attributes like caste in India or household registration in China, due to its impact on individuals' economic well-being, social inequality and inter-generational mobility (Greenwood et al. (2014), Gihleb and Lang (2020), Banerjee et al. (2013), Han et al. (2015), Almar et al. (2023)). Most of the previous work focuses on the identification of existence and intensity of assortative matching and how it affects socioeconomic outcomes. In this paper, based on the traditional roles assumed by men and women in a household, I account for the difference between the spouses in the preferences alongside the assortative matching.

Two heavily studied mating preferences in the literature do not fit the patterns in the data well theoretically. Positive assortative matching is one of the well-documented and measured preferences. By assuming positive assortative matching and that education is a valuable trait in the marriage market, the least educated group of the gender with a population in surplus would be forced to be single. For example, in China, overall there are more men than women. In this setup, the least educated men should have a positive single rate, while all other groups are predicted to be married. The matching should remain stable even if the education distribution of women shifts up or down because it is not the absolute amount of education but the rank of an individual in the education distribution of his or her gender matters. In real data, we indeed observe a high single rate for men with least education, but the single rates for other education groups of men are not negligible, too. It also fails to rationalize all the single rates of women.

Other than positive assortative matching, homogamy is another candidate to consider. Both women and men prefer to marry someone with similar education and get a disutility from both marrying up and down. It fits the marriage patterns reasonably well (Figure 7 and 8). Couples with matching education are most common, although they are a few groups for whom the degree of homogamy is less strong. However, given the relative sizes of each education and gender group in our sample based on the Census, and if we assume people would rather stay single if not married to a spouse with the same education, homogamy would predict single rates inconsistent with what we observe in the data, as shown in Figure 2. The relationships between single rates and education for two genders are almost the opposite to the patterns in the Census. It reflects the surplus of available females and males at each education level in the marriage market. If the distaste for unequal education in the marriage is not so extreme, it is not obvious how the single rates would be like and an empirical estimation of preferences would be necessary.

2 Data

I use the 1% sample from the 2010 Census in the analysis of this paper, the most recent Census currently available.

To detect patterns in the overall marriage market, my initial sample consists of individuals born between 1960 and 1975. I drop a small number of who were enrolled in school at the time of the survey. The restriction on birth cohorts are imposed because I am interested in the relatively young population in the sample whose educational and marital outcomes have most likely unfolded.

Figure 1, 3, 4, 5 and 6 are based on this sample.

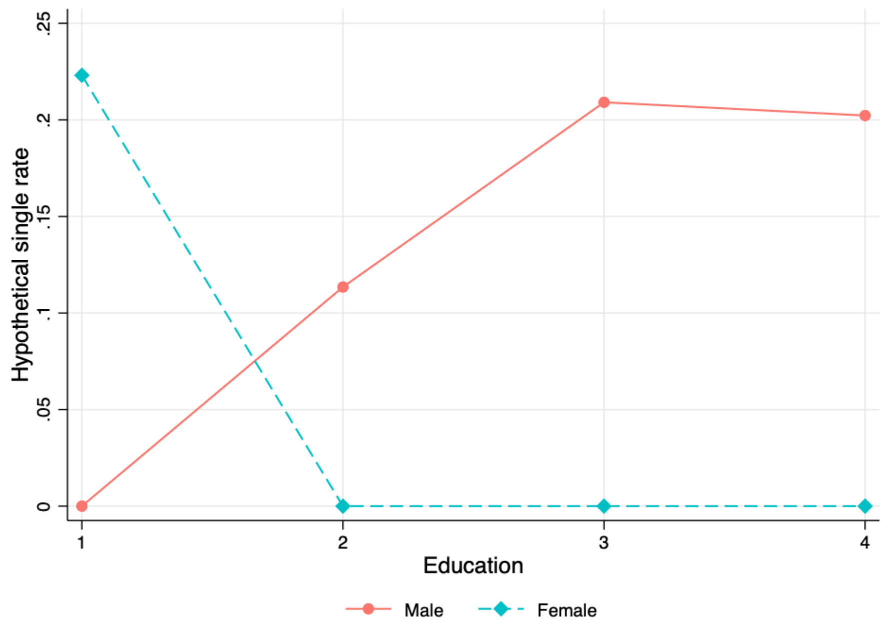
2.1 Several stylized facts

2.1.1 Single rates

A new definition of being single versus being married

Before discussing the patterns in single rates, it is important to clarify that the definition of being single used in this paper differs from the common one adopted by the Census. Here, single rates are defined as the ratio of individuals who are single at the time of

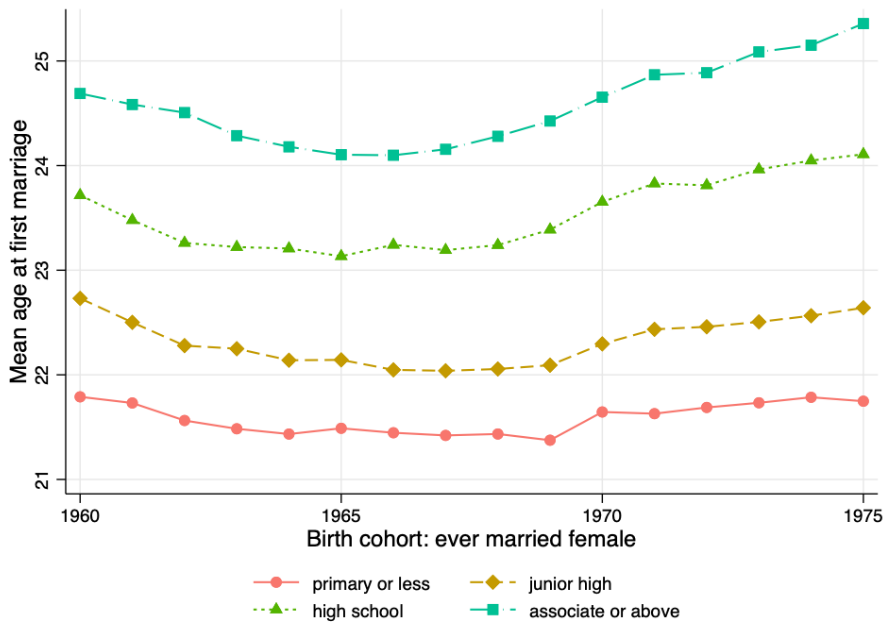
Figure 2: Hypothetical single rates with the extreme homogamy



Note: This figure depicts the hypothetical single rates for females and males who were born in 1975, i.e. 35 years old at the time of survey by education level and birth cohort. All individuals born in 1975, not in school at the time of survey, and were married, divorced, or widowed with a non-missing value for marital status are included. Those who married for the first time after 35 are treated as single in the calculation. The extreme homogamy is defined to be the case where people marry spouses with equal education or stay single. Data source: Census 2010 1% sample.

the Census or who marry for the first time after the age of 35 to the total population, categorized by gender and education level. The cutoff age of 35 was chosen because the majority of people marry by that age. In the sample, only 0.3 % of the currently married females and 0.9 % of the currently married males married after 35. For those who married for the first time before 35, the mean ages at first marriage are summarized in Figures 3 and 4. The mean age slightly rises for the two more educated groups for both women and men, while it remains relatively constant for the two less educated groups. For individuals born in 1975, the mean age at first marriage is around 25.3 for women and 27 for men, which is significantly younger than 35. This further corroborates that using 35 as a cutoff is not too restrictive. Setting this cutoff ensures an equal time horizon for marriage across older and younger cohorts; otherwise, the single rates for the younger cohorts would be mechanically biased up. Matched fractions are defined as the ratio of the matched population in an education pair over the the total population, given the gender and education level. By definition, the single rate and matching fractions for a specific gender and education level add up to 1. Any mention of being single or married throughout this paper refers to these customized definitions, unless noted otherwise.

Figure 3: Mean age at first marriage for females

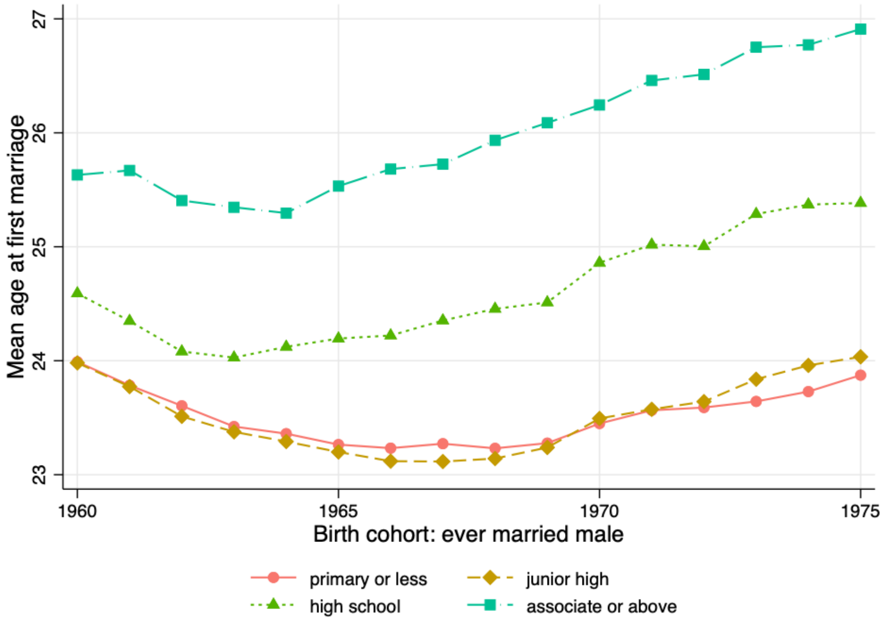


Note: This figure summarizes the mean ages at the first marriage for females who were at least 35 years old and no longer in school at the time of survey by education level and birth cohort. All women who were married, divorced, and widowed with a non-missing value for the age at the first marriage are included. Ages at the first marriage above 35 are not used for calculation. Data source: Census 2010 1% sample.

Two key features of interest

The first feature of single rates that this paper aims to explain is the contrasting variation in single rates by education level between women and men. Specifically, I focus on the single rates for the most recent birth cohort of 1975 (Figure 1). In general, single rates increase with education for women, although they are nearly identical for those with junior high school education and the least educated. Conversely, single rates decrease overall as men attain higher levels of education. Moving from the junior high school group to the high school group does not significantly reduce the likelihood of a man remaining single by age 35, though. It is not immediately clear what kinds of preferences could explain this positive relationship between single rates and education for women,

Figure 4: Mean age at first marriage for males



Note: This figure summarizes the mean ages at the first marriage for males who were were at least 35 years old and no longer in school at the time of survey by education level and birth cohort. All men who were married, divorced, and widowed with a non-missing value for the age at the first marriage are included. Ages at the first marriage above 35 are not used for calculation. Data source: Census 2010 1% sample.

and the negative relationship for men.

Furthermore, single rates have increased across all groups among younger cohorts. Notably, both women and men exhibit a widening gap in single rates between the more educated and the less educated within these cohorts (Figure 5 and Figure 6). Given the specific definition of being single used in this study, this widening gap is unlikely to be solely due to more educated women and less educated men delaying marriage. Instead, it reflects significant changes in their decisions regarding whether to marry at all. Understanding the role of mating preferences behind these trends sheds light on ongoing discussions around phenomena such as the "left-behind women" and may help destigmatize certain population groups.

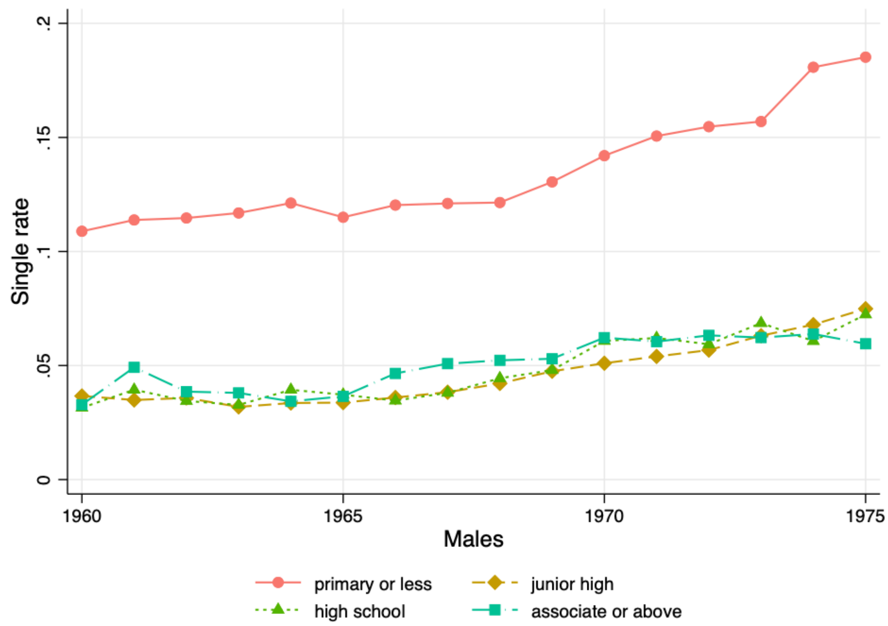
Figure 5: Single rates for females

Note: This figure summarizes the the single rates for females who were at least 35 years old and no longer in school at the time of survey by education level and birth cohort. Those who married for the first time after 35 are treated as single in the calculation. Data source: Census 2010 1% sample.

2.1.2 Marriage matching patterns

Homogamy, a matching phenomenon documented by researchers in many contexts, exists among people born between 1972 and 1975 in China. People are more likely to marry someone with approximately the same education as themselves have. However, they do not always find a spouse with equal education. In fact, the assumption on the aversion to a marriage in which the wife is more educated is partly motivated by the difference between men's and women's matching patterns if they do not marry someone with equal education. In Figure 7 and 8, we can see that among those with junior school and high school education, for whom both marrying up and down are possible in the categorization of this paper, aside from matches within the same education level, men are more likely to marry down rather than up, while women exhibit similar likelihoods of marrying up and down, slightly being more likely to marry up. Although the least educated group are not possible to marry down by the definition of education groups I use, the pattern on the other side that the number of the least educated men marrying slightly more educated women is far outnumbered by the the number of them marrying women as least educated

Figure 6: Single rates for males



Note: This figure summarizes the single rates for males who were at least 35 years old at the time of survey by education level and birth cohort. All men who were married, divorced, and widowed with a non-missing value for the age at the first marriage are included. Those who married for the first time after 35 are treated as single in the calculation. Data source: Census 2010 1% sample.

as them also supports this hypothesis of aversion. Similarly, this aversion could explain why the gap between the numbers of the most educated women marrying men with equal education and men with slightly less education is much larger than that in men's case.

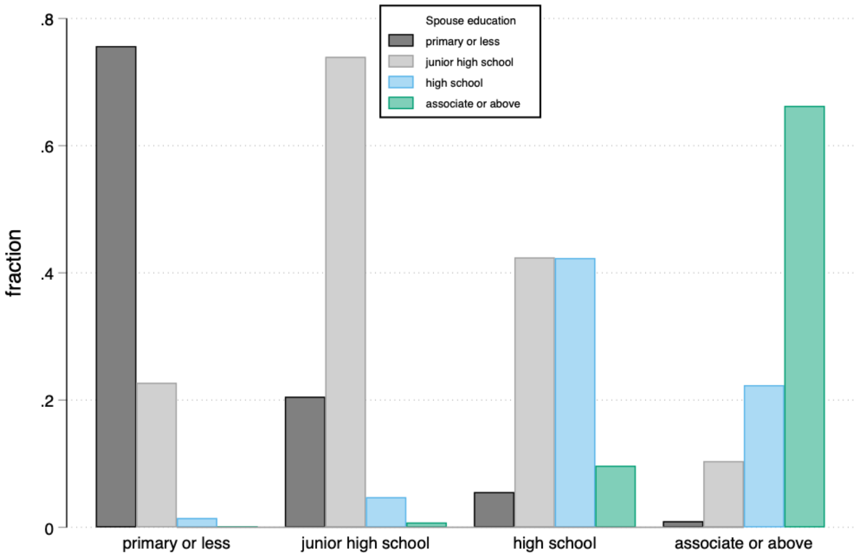
2.2 Sample construction

For the purpose of estimating preference parameters in the utility model, the simulated sample is constructed in the way I describe in detail below. Figure 7 and 8 are based on this simulated sample for estimation.

To calculate matching patterns, I first identify married couples or single headed households from the sample. Most households can be identified by specifying household ID and residential space features, like the home area and addresses. In some cases, multiple single individuals share all these information and they are all household heads. I treat

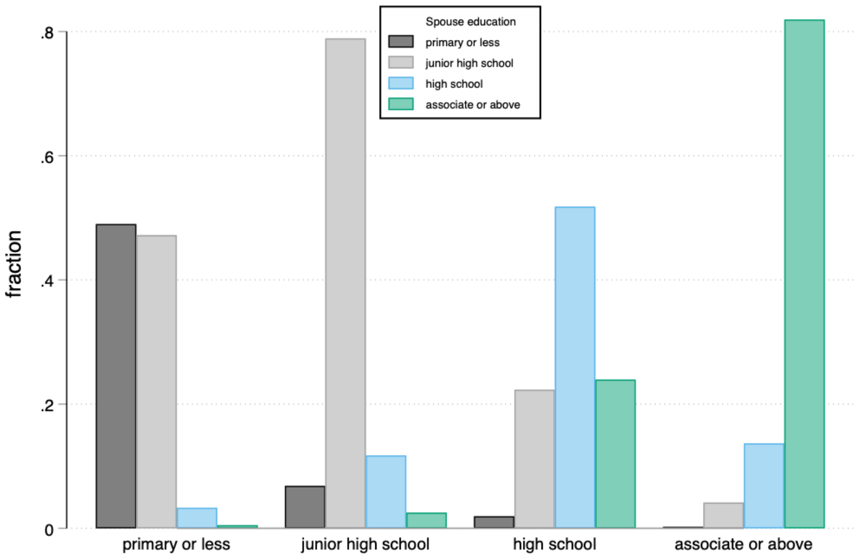
[1]Refer to Section 2.2 for more details.
[2]Refer to Section 2.2 for more details.

Figure 7: Marriage matching pattern: males



Note: This bar plot summarizes the distributions of spouse (female) education among married men with one of the couple born between 1972 and 1975.[1] Data source: Census 2010 1% sample.

Figure 8: Marriage matching pattern: females



Note: This bar plot summarizes the distributions of spouse (male) education among married women with one of the couple born between 1972 and 1975.[2] Data source: Census 2010 1% sample.

them as separate households. One example is the migrant workers in cities sharing an apartment. I only keep the individuals who are household heads themselves or the spouses of the household heads because the marriage relationships are harder to be pin-pointed among the rest of the household relationships relative to the the head in my data set. Then I further drop couples with the different first marriage years or months. The data set does not have direct information on whether the current marriage is the first one for each of them. Different first-marriage times indicate that at least one of the couple had one or more previous marriages. The dynamics for a setup with re-matching could be more complicated and I abstract away from that in my analysis. Besides, the fraction of married couples with different first marriage times is quite small. Among all the households with one married couple, only 2.2% of them fall into this category. The number is close to the divorce rate[3], 2.13 %, defined in a similar fashion around 2010. Ignoring this group should not threat the generality of the story. In the cases where there are more than one set of household heads and spouses, I separate the couples by matching males and females as couples if they have the same year and month for their first marriages. This could happen when siblings and their spouses live in the same house, which are not common overall and mostly exist in rural areas. If a couple have different times for the first marriage, they are also dropped in this case.

Due to the simplified attributes of the matching process, a woman or man can either be single (unmatched) or married (matched). There is no scope for being divorced or widowed. Moreover, the data does not information on ex-spouses including education and age for the divorced and widowed. Hence, when calculating the population moments, I exclude the divorced and widowed from the population counts as well.

In the constructed sample, the numbers of married females and males must be equal. Therefore, in the Census, if one person's marriage age is below 35 while his or her spouse is married after the age of 35, I consider both of them to be married after 35, thus being categorized as single. I restrict the Census sample to consist of people at the age of 35 or above, because the inclusion of younger cohorts might inflate the single rates and bring down the mean age at the first marriage. Thus, the youngest cohort from the Census would be born in 1975.

In the following estimation, I pool together 4 cohorts born between 1972 and 1975. It reduces the year by year fluctuation in the marriage market and also increases the sample size, which could be particularly useful in the estimation for the highest and lowest education categories where there are less people. Again, to avoid the situation

[3]https://www.globaltimes.cn/page/201809/1120041.shtml

where one side of the couple is dropped because he or she was not born in this range of years while his or her spouse was or vice versa, I keep the couple as long as one of them was born in the selected birth cohorts. Since it is not uncommon for people to marry someone a few yeas younger or older, a significant amount of individuals born in the years close to the range of 1972 and 1975 are kept in the sample because their spouses were born in the range. Figure 9 shows the distributions of birth cohorts for the extended sample born before 1972 by gender. There are more males distributed on this side. This is due to that women are more often to marry someone older than men do. Figure 10 contains the distribution of birth cohorts after 1975 by gender. There are more females distributed on this younger side due to the same reason.

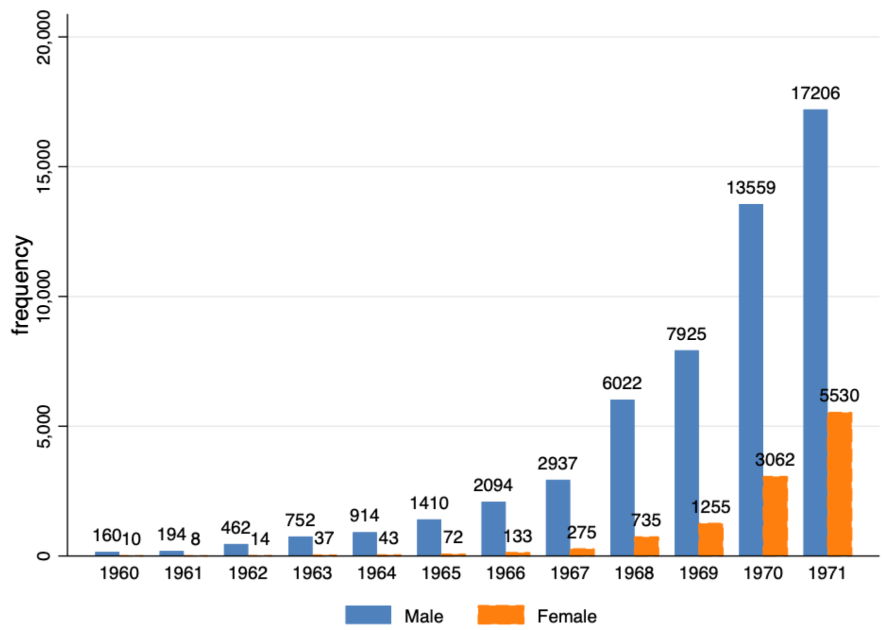
One caveat of this practice is that all the additional individuals included in the sample, who were not born within the targeted age range, are married, which will bias the single rates downward. To address this, I construct the simulated sample in the following manner. The proportion unmarried in each gender/education group is the proportion in the data for the 1975 cohort. The distribution of matches across the remainder is the education match distribution for the expanded sample including spouses born outside the 1972-75 cohort. See Figure 18 in Appendix for a simplified graphic illustration of this process.

3 Model

Utility. In this section, I model the utility of men and women from marriage and remaining single. To simplify the estimation, the utility of being single is set to be 0 for both males and females.

I assume that a man's utility from marriage does not only depend on the wife's education level but also the difference in education levels of the couple. Specifically, if the wife's education is higher, the man cares about the difference between their educations, although I do not impose assumptions on whether women in such case become more or less desirable in the marriage. A man m's utility of marrying a potential spouse, woman f, relative to remaining single, is given by Equation 1.

Figure 9: Spouse's age not in the targeted birth cohorts: Born before 1972



Note: This bar plot summarizes the distributions of birth cohorts for spouses born before 1972. They are included in the sample because their spouses were born between 1972 and 1975.Data source: Census 2010 1% sample.

$$u_m = a_f + b \cdot educ_f + c \cdot \max(educ_f - educ_m, 0) + d \cdot \max(educ_f - educ_m, 0)^2 + k \cdot \mathbb{1}(educ_f - educ_m > 0) + \epsilon_{mf} \tag{1}$$

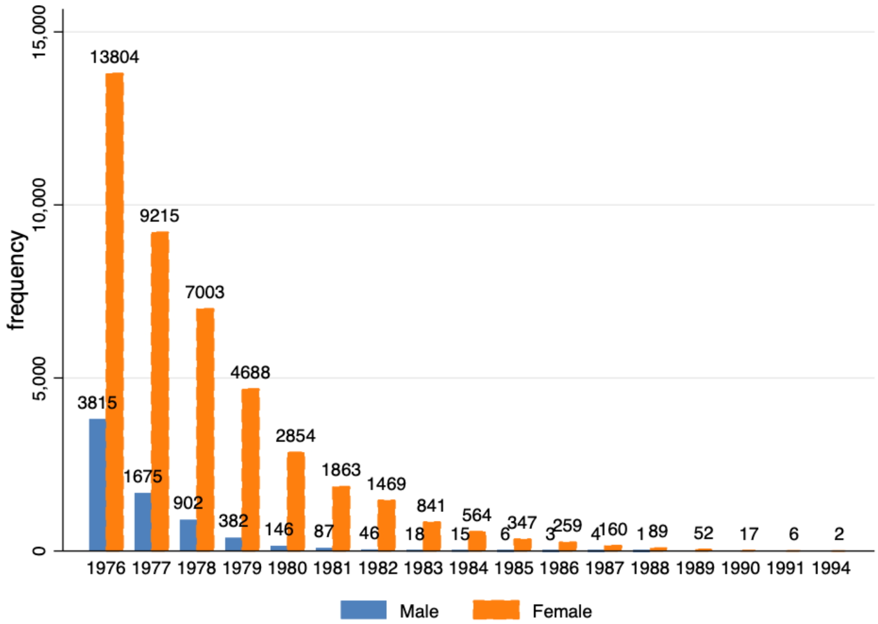
ϵ_{mf} is the random taste of the man m for women f. It follows a normal distribution of mean 0 and standard deviation 5. The choice of standard deviation only affects the magnitude of the estimated coefficients but not how well the matching outcome is.

a_m is a constant that varies by men's education. I use 4 education categories when estimating the preferences: primary school or below including no schooling at all, junior high, high school, associate degree or above, including four-year college, and graduate school.

The utility of marriage relative to being single for females is modeled in a similar fashion.

$$u_f = \alpha_f + \beta \cdot educ_m + \gamma \cdot \max(educ_f - educ_m, 0) + \delta \cdot \max(educ_f - educ_m, 0)^2 + \kappa \cdot \mathbb{1}(educ_f - educ_m > 0) + \epsilon_{fm} \tag{2}$$

Figure 10: Spouse's age not in the targeted birth cohorts: Born after 1975



Note: This bar plot summarizes the the distributions of birth cohorts for spouses born after 1975. They are included in the sample because their spouses were born between 1972 and 1975. Data source: Census 2010 1% sample.

Matching. I abstract away from information friction in the matching procedure and assume the utility is non-transferable. I adopt the deferred acceptance algorithm pioneered by Gale and Shapley (1962) to perform the matching. Men (m_1, m_2, ..., m_M) and women ((w_1, w_2, ..., w_W) rank all the individuals on the other side based on the utility they can get if matched/married to that individual, given by Equation 1 and 2 above. Because I have set the utility of being single to 0, a spouse would be considered as unacceptable if the the utility of marrying him or her is less than 0. One side of the agents should propose while the other side make decisions on whether to reject or keep the proposal. I assume men propose, which is the social convention in China. The matching starts. In the first round of proposing, all the men propose to their first choices in the rank lists. If his set of acceptable spouses is empty, he does not propose to anyone and will remain single in the end. Each woman could receive zero or a positive number of proposals. She holds the most preferred acceptable proposal and rejects all the others. If all the proposals she receives are from her unacceptable spouse set, she rejects those proposals. When the second round starts, if a man 's proposal gets rejected in the first round, he moves down on his preference list and proposes to the second preferred mate among the acceptable ones. If his proposal in the first round is on hold, he does not propose to a new woman in this round. After the proposals in the second round are made, women update the best proposal received and reject the rest. This process repeats until no proposals are made. Men with proposals not rejected at this stage get matched to the women who are holding their proposals. The rest remain single or unmatched. The steps of the matching procedure is summarized in Table 1.

Although the deferred acceptance algorithm is highly stylized, it captures some essential features of the dating and marriage process. Men reach out to their most favored woman first and, if rejected, move on to their second choice. Women date the best man who has proposed to them so far. If they get a proposal from a man they like more, they reject the one that they are currently dating. Women get settled as being single or married to the partner they are currently dating when no more proposals come in. Men get settled when the women that they like the most, among the ones that have not rejected them yet, decide to marry them or stay single after exhausting their choice set before finding a stable match. Some alternative models, like sequential search models or Becker's models with transferable utilities, also abstract away from some details of the dating and marriage and do not present themselves as superior to the deferred acceptance algorithm in an economic or computational sense.

Modern technology makes it common for agents to submit a preference list and implement the matching in a centralized system, as we see in its application in a student college matching problem. However, a centralized system is not necessary for carrying out a deferred acceptance style matching process. The decentralized marriage matching process described above also has the key features of a deferred acceptance algorithm.

Algorithm 1 Matching in the Marriage Market: Deferred Acceptance Algorithm (Gale and Shapley, 1962)

Input: Men $\{m_1, m_2, \ldots, m_M\}$ and Women $\{w_1, w_2, \ldots, w_W\}$, each with ranked preferences.

Output: A stable matching between men and women.

Step 1: Each person ranks all individuals on the other side based on the utility they can get if matched/married to that individual.

Step 2: Each person has a utility of being unmatched (single).

Step 3: One side of the agents proposes (e.g., men propose):

- **Step 3a:** Each man proposes to his most preferred acceptable choice.
- **Step 3b:** Each woman holds the most preferred and acceptable proposal she received and rejects all the others.

Step 4: The process repeats until all agents are matched or no more acceptable proposals can be made.

4 Estimation

I calculate the marriage market moments based on the sample constructed in the way described in Section 2.2. Then I create the sample to estimate the parameters in the utility functions by scaling down the numbers of men and women in each education category. The loss function is constructed as the sum of squared deviations in the numbers in each matched category and single category.

The loss function is defined as in Equation 3. The education category set is the same for both men and women, which is E = {primary or less, junior high, high school, associate or above }. K is the set of education pairs for matched couples. K = E ⊗ E.

$$L(\hat{\theta}) = \underbrace{\sum_{(m,f)\in\mathbf{K}} (g_m(\hat{\theta}) - g_{m0})^2}_{\text{male matched Ns}} + \underbrace{\sum_{m\in\mathbf{E}} (h_m(\hat{\theta}) - h_{m0})^2}_{\text{male single Ns}} + \underbrace{\sum_{(m,f)\in\mathbf{K}} (g_w(\hat{\theta}) - g_{w0})^2}_{\text{female matched Ns}} + \underbrace{\sum_{w\in\mathbf{E}} (h_w(\hat{\theta}) - h_{w0})^2}_{\text{female single Ns}} \tag{3}$$

The Newton-Raphson method is not applicable in this case due to the non-differentiable and non-convex nature of the matching process. The Nelder-Mead method is employed instead. To reduce the chance that the optimization leads to local minima, I optimize over as many combinations of parameters as possible, starting from the one-parameter combinations. Table 1 summarizes the numbers for each education and gender type in the calibration sample. The estimates of parameters are summarized below in Table 4. The prediction precisions are shown in Table 2, Table 3 and Figure 11.

Table 1: Education distribution in the calibration sample

Education	Male	Female
primary or less	295	379
junior high	921	817
high school	239	189
associate or above	206	165
Total	1661	1550

Discussion The prediction fits the single patterns in the calibration sample relatively well. The predicted single rates reproduce the most distinctive features in the single rates we observed in the Census 2010, where women's single rates roughly increase in education while men's single rates roughly decrease in education (Figure 11). However, in the prediction, these relationships are further strengthened. In particular, in the data, the single rate for women with primary education or less is not much different from that for women with junior high school education. In the prediction, the single rate for the primary or less is under-predicted to be 0. Similar prediction errors apply to men with associate education or more. Other than these two groups, single rates for the rest are closely fitted. [4] The marriage matching patterns are also closely fitted. The results are summarized in Table 2 and Table 3). The degree of positive assortative matching is well approximated. Precisely, the prediction retains two patterns in the data: A vast majority in each education and gender type marry people with the same level of education; other than marrying people with the same education, women are more likely to marry men with more education than themselves and men behave oppositely.

When interpreting the estimated parameters in Table 4, their ordinal properties and signs matter the most. Both men and women prefer a less educated spouse, although men's preference is much stronger (b = -4.56 for men, β = -1.51 for women). This set of preferences differs from the usual expectation that everyone prefers the most educated.

[4]The loss function in Equation 3 takes a value of 559 in terms of the sum of squared deviations in numbers of people or 0.0099 in terms of the sum of squared deviations in fractions.

Still, the preference for the less educated predicts the positive assortative matching pattern in the data. In addition to the preference linear in spouses' education, men have a solid aversion to marrying women more educated than themselves (c, d, e for men; γ, δ, ϵ for women). If the wife receives one more unit of education than the man, the man's aversion is measured by (-7.92 -0.79 + 0.08). Women do not like less educated men, either. However, the aversion is much less intense. In the example above, a woman's aversion is measured by (-0.80 + 0.47 -2.23). Figure 6 plots the mean utilities of men and women at each education level from marrying spouses with different education. It gives us a general sense of the mean utility of getting married overall for all education and gender groups.[5] Overall, the most educated men get the highest values out of marriage while the least educated men have the lowest evaluations for marriage. Among women, the least educated have the highest value of marriage, followed by the most educated. Under the traditional social norm, men are the breadwinners of the family, and women are the housekeepers. This stereotype of gender roles could explain the overall evaluations for marriage for the most and the least educated men and the least educated women. For the most educated men, the relatively high incomes allow them to provide for the family without too much stress. However, for the least educated men, providing for the family may exacerbate their economic conditions so much that the disutility from worse financial conditions dominates the extra benefits they could get from marriage, resulting in a low value of getting married. The least educated women typically work in occupations that are intensive in manual tasks. They often earn less than men in those jobs. Hence, marriage may improve their economic conditions. It is unexpected to see that the most educated women, who are more likely to be economically independent, also have a high value of marriage. One potential reason could be that these groups of women attach high value to some attributes of the marriage, such as passing human capital to the next generation. However, more research is needed on the reasons behind their attitudes towards marriage.

Circling back to the question I raised in the beginning that why the highly educated women have the highest single rates among women of all education levels, the results suggest it is mainly due to men's aversion to marrying a more educated spouse. The

[5]The constant parameters a_e and α_e could not be directly interpreted as the mean utility of an average marriage, although they are factored into the values of marriage. If we look at one of the eight constants, each corresponding to one education and gender group, the spouse education term and the terms related to the education differences cannot be 0 simultaneously

mean utilities of the highly educated women from the marriage are relatively high. It means that many women in this group do not stay single by choice. Men's preference for less educated spouses, compounded by their strong aversion to marrying a woman more educated than themselves, make the most educated woman least preferred among men. Men with primary education or less, junior high school education, and high school education have a negative mean utility from marrying the most educated women, which means after taking into account the random taste, a man in these education groups would rather stay single than marry more than half of these well-educated women. Men with associate education or above are the group that derives the highest utility from marrying them. The mean utility, 20, is about 4 standard deviations above 0. [6] Thus, this group of well-educated men would almost surely prefer marrying an equally well-educated woman over staying single. However, the most educated women are still ranked last on their preference list, which means most of them would only make proposals to the most educated women if their proposals to the less educated women are turned down in the earlier rounds of the matching process.

When it comes down to the high single rates of another group, the least educated men, in contrast, the results suggest that it is more of a choice. They are very willing to marry women also with the least education. However, their mean utilities from marriage decrease sharply as the spouse's education increases, resulting in an overall low utility from marriage. As Figure 12a shows, their mean utility of marriage turns negative if their spouse's education increases from primary school or less to junior high school. Their mean utility reaches more than four standard deviations below the utility of being single, 0, when considering a woman with associate education or above.

Notably, although women prefer less educated spouses like men, this preference is less intense. Unlike the preferences of the least educated men discussed above, women's preferences for less educated spouses are further weakened by their aversion to marrying a man less educated than themselves. For example, the least educated men are not ubiquitously perceived as the favorite by women, while the least educated women dominate women at other education levels in all men's evaluations. They almost rank the same as high school-educated men in the standard of high school-educated women. Junior high school-educated women prefer men with the same education over these least-educated

[6] Recall that the random tastes follow a normal distribution of mean 0 and standard deviation 5.

men. [7] Men with high school education are almost equally favored as the least educated men by this group of women. Surprisingly, the sex ratio imbalance is not a direct reason for the highest single rate of the least educated men. There are more women with primary or less education than men, although men outnumber women in total (See Table 1).

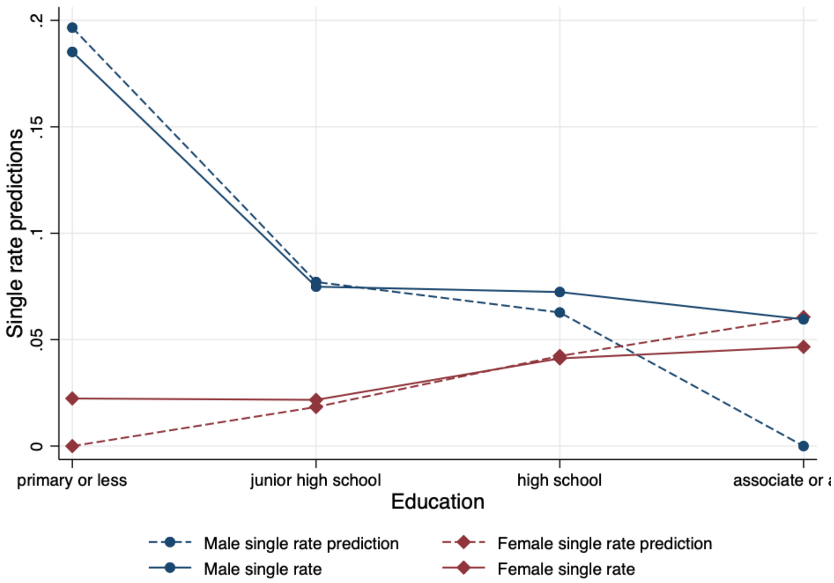
The other two groups with the lowest single rates are not the focus of the research question, but still, the results provide a straightforward explanation of the underlying preferences. The least educated women are most favored on average due to men's strong preferences for a less educated spouse. This group of women would get most of the proposals in the early rounds. Although the distaste for spouses' education could justify the low single rates for the least educated women among other education groups, further model extensions are needed because it currently fails to predict any single women in this education group.

The most educated men are also predicted to all get married. They are not most favored by any group of women. This result mainly reflects their high evaluation of marriage overall, and thus, more of them have a larger set of acceptable spouses. In the meantime, unlike highly educated women, they are not hurt much by having more education, given that women's mean utilities are overall flat in the spouse's education.

[7] Although the mean utility of women with junior high school education marrying men at any education level would be less than that of being single, each woman in this group still has a non-empty set of acceptable men due to the random tastes drawn for all potential mates every woman faces. Generally, the mean utility being negative does not necessarily mean the majority of this group being single.

[8] Numbers are shown in Table 6

Figure 11: Single rate prediction



Note: This figure summarizes the prediction of single rates and the true single rates in the sample for the estimation.

Table 2: Prediction precision- male matchings

Category	Target	Prediction	Deviation
primary or less: primary or less	181.49	180	-1.49
primary or less: junior high	54.51	53	-1.51
primary or less: high school	3.49	4	0.51
primary or less: associate or above	0.37	0	-0.37
primary or less: single	54.51	58	3.49
junior high: primary or less	174.84	174	-0.84
junior high: junior high	630.03	629	-1.03
junior high: high school	40.32	41	0.68
junior high: associate or above	6.48	6	-0.48
junior high: single	68.95	71	2.05
high school: primary or less	12.23	18	5.77
high school: junior high	93.75	95	1.25
high school: high school	93.56	95	1.44
high school: associate or above	21.4	16	-5.4
high school: single	17.24	15	-2.24
associate or above: primary or less	1.86	7	5.14
associate or above: junior high	20.14	25	4.86
associate or above: high school	43.25	41	-2.25
associate or above: associate or above	128.19	133	4.81
associate or above: single	12.24	0	-12.24

Note: All the numbers are rounded to the second decimal place. The estimation sample takes the sample with matched couples from the Census and scale down by 1/100 to keep time needed for matching in a reasonable range. This is the reason why the targets are not integers. Deviations less than 1 can be ignored and perfect predictions have been achieved. For row "primary or less: junior high", the entry stands for the number of males with primary or less education marrying females with junior high school education, as the target or as the predicted. Similar definitions apply to the rest of the table.

Table 3: Prediction precision- female matchings

Category	Target	Prediction	Deviation
primary or less: primary or less	181.49	180	-1.49
primary or less: junior high	174.84	174	-0.84
primary or less: high school	12.23	18	5.77
primary or less: associate or above	1.86	7	5.14
primary or less: single	8.47	0	-8.47
junior high: primary or less	54.51	53	-1.51
junior high: junior high	630.03	629	-1.030
junior high: high school	93.75	95	1.25
junior high: associate or above	20.14	25	4.86
junior high: single	17.72	15	-2.72
high school: primary or less	3.49	4	0.51
high school: junior high	40.32	41	0.68
high school: high school	93.56	95	1.44
high school: associate or above	43.25	41	-2.25
high school: single	7.75	8	0.25
associate or above: primary or less	0.37	0	-0.37
associate or above: junior high	6.48	6	-0.48
associate or above: high school	21.4	16	-5.4
associate or above: associate or above	128.19	133	4.81
associate or above: single	7.64	10	2.36

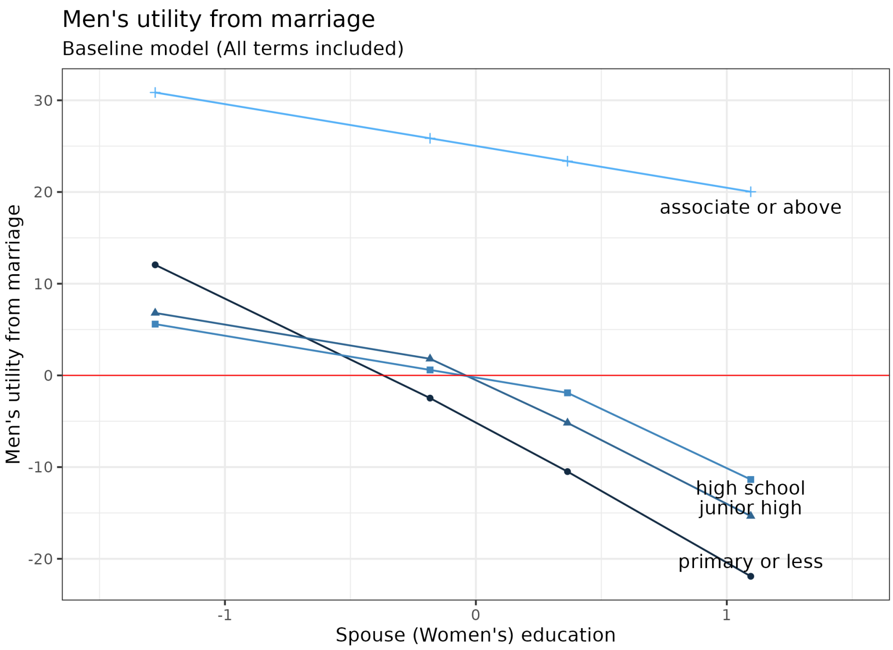
Note: All the numbers are rounded to the second decimal place. The estimation sample takes the sample with matched couples from the Census and scale down by 1/100 to keep time needed for matching in a reasonable range. This is the reason why the targets are not integers. Deviations less than 1 can be ignored and perfect predictions have been achieved. For row "primary or less: junior high", the entry stands for the number of females with primary or less education marrying males with junior high school education, as the target or as the predicted. Similar definitions apply to the rest of the table.

Table 4: Estimated preference parameters

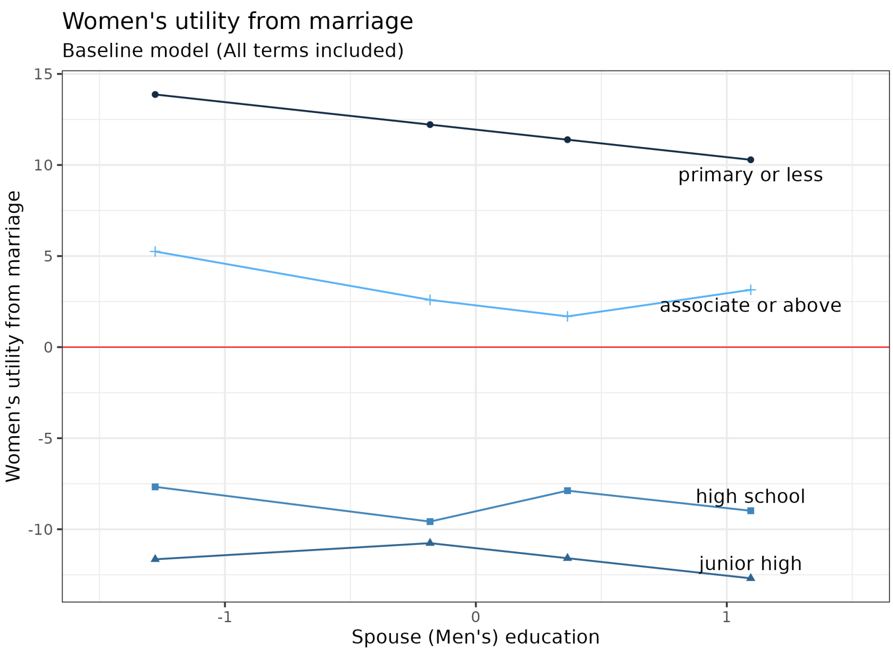
	estimates			
Male				
a_e	6.23	0.98	-0.24	25.02
b	-4.56			
c	-7.92			
d	-0.79			
e	0.08			
Female				
α_e	11.94	-11.04	-7.33	4.80
β	-1.51			
γ	-0.80			
δ	0.47			
ϵ	-2.23			

Note: All the estimates are rounded to the second decimal place. a_e and α_e have four values that correspond to males and females in four education categories, respectively.

Figure 12: Mean utilities from marriage [8]



(a) Baseline: men's utility from marriage



(b) Baseline: women's utility from marriage

5 Single rates when men's or women's aversion is turned off

As discussed in the previous section, men's distaste for the most educated women can be decomposed into two parts. First, their utility from marriage decreases as the spouse's education increases. It comes from the negative coefficient, b, of the spouse education term in their utility function. Second, there is an additional loss in their utility if the spouse has more education than they do. This is summarized by a negative effect jointly from the coefficients, c, d, and e when a man's spouse has higher education. Similarly, women's utility from marriage consists of two parts. In this counterfactual analysis, I keep the coefficient for the linear term of the spouse's education unchanged. I examine the effects on single rate patterns, especially for the most educated women and the least educated men, when the coefficients for terms relating to the differences between education levels of the couple are reset to zero for men and women, respectively. More straightforwardly, I investigate single rates when c, d, and e are set to be zeros, or γ, δ, and κ are set to zeros, respectively.

When men's aversion to a more educated wife is turned off in their utility function, men with primary education or less are affected the most. Their utility of marrying women with education more than themselves increases significantly (Figure 15a). In the baseline model, an important reason why many men in this category stay single is because they have a strong aversion to marrying women with more education. Blocking the distaste for a more educated wife is shown in Figure 13 to be powerful enough to reduce the single rate for this group of men to zero. In the meantime, the predicted single rates for women also drop to a near-zero level. Many more women are matched to the least educated men whose preferences change the most in this scenario. Noticeably, the single rate for high school-educated men spikes. This group of men now becomes the group that has the lowest utility from marriage overall. Compared to the baseline model, this framework fails to predict the high single rates of the least educated men and most educated women. See Figure 13 for predicted single rates and Figure 15a for predicted mean utilities from marriage for men in this counterfactual case. Table 9 with predicted mean utilities for both genders, Table 7 and Table 8 with prediction precisions are in the Appendix.

Alternatively, I set the women's utility function to be linear in their spouse's education and mute women's aversion to marrying a less educated man. In this case, the single rate for the least educated men still drops compared to that in the baseline model, but the magnitude is smaller compared to the case where men's aversion is turned off. The change in women's single rates is similar to the first case. Except for the junior high school-educated women, other groups all have a zero or near-zero single rate. See Figure 14 for predicted single rates, Figure 15b for predicted mean utilities from marriage for women, and Table 10 and Table 11 for prediction precisions in this counterfactual case.

Turning down the aversion terms in the first counterfactual case affects the utility of the least educated men so much that it not only makes the mean utility from marrying a more educated wife positive, which means many more of the least educated men would start making proposals instead of choosing to be single but also lifts their values of marrying women with junior high school education, high school education and associate education or above ahead of junior high school-educated and high school-educated men. This means that a higher fraction of them, compared to junior high school-educated and high school-educated men, would make proposals to these women if they were not matched to the least educated women.

On the contrary, in the second case, men's aversion to a more educated spouse is still in place, but women's aversion to a less educated spouse is removed. The magnitude of change in women's mean utilities from marriage is not as large. Specifically, for the three lower education groups of women who are affected, the change neither flips the signs of the mean utilities from marriage nor alters the relative positions of their mean utility lines. This means that most women would not change their decision between getting married and staying single. While their chances of accepting a marriage proposal from a man with a certain level of education may have increased, the order of those chances remains the same. As a result, most men go through a similar process of making proposals. Still, the utilities from marrying a man with the least education increase for women with junior high, high school, and associate degrees or higher, so proposals made by these men are more likely to be accepted. Hence, we see that the single rate for the least educated men still decreases even though their own preferences do not change.

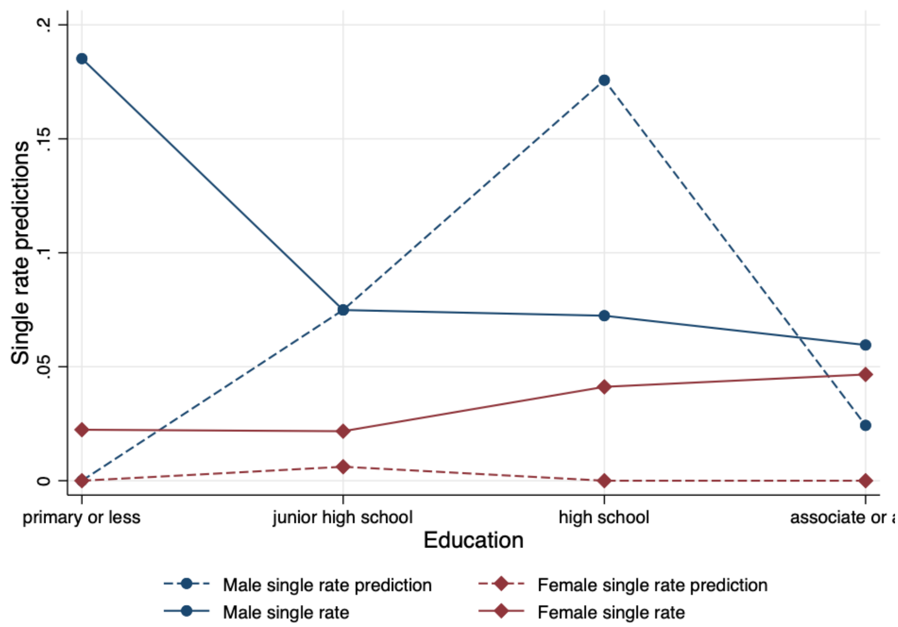
It is worth pointing out that although the single rate for women with the highest education can be reduced to zero in both counterfactual simulations, their implications on the matching outcome are different. Table 5 compares the marriage and single patterns of the most educated women in these two cases. When women's aversion to a less educated spouse is turned off in Simulation 2, most women who would have been single in the baseline prediction are now predicted to be married. They are mainly matched to equally highly educated men with associate education or above or high school-educated men. In Figure 12, we observe that other than the least educated women, women's value of marrying a man in the three lower education groups increases overall. Thus, the proposals from men in these three groups become more likely to be accepted/kept than in the baseline case because women like marrying them more. The most educated men do not benefit from the changing preference of women in this counterfactual world. They become more likely to be rejected in the earlier rounds of proposing to less educated women than in the baseline model. More of them propose to equally educated women in the later rounds. This is why we observe that the number of highly educated women matched to this group of highly educated men increases the most. A similar logic explains why the number of highly educated women matched to high school-educated men increases the second most. All three higher educated groups increase the value of marrying a least educated man; the least educated men benefit the most. High school-educated men benefit only from the increased evaluation of women with associate education or above. Their desirability does not change when proposing to the less educated women in the earlier rounds. In fact, their relative competency among junior school-educated and high school-educated women decreases because these women now view the least educated men and junior high school-educated men more favorably but do not change attitudes towards the high school-educated men. However, in Simulation 1 where men's aversion to a less educated spouse is turned off, significantly more women marry the two less educated groups of men. Not only those who stay single before marry them, but also those who marry the high school-educated men and men with associate education or more. The positive assortative matching pattern is no longer preserved.

Table 5: Marriage/single patterns of the most educated female with different preferences

Category	Target	Baseline	Simulation 1	Simulation 2
associate or above: primary or less	0.37	0	43.00	0.00
associate or above: junior high	6.48	6	55.00	6.00
associate or above: high school	21.40	16	11.00	18.00
associate or above: associate or above	128.19	133	56.00	141.00
associate or above: single	7.64	10	0.00	0.00

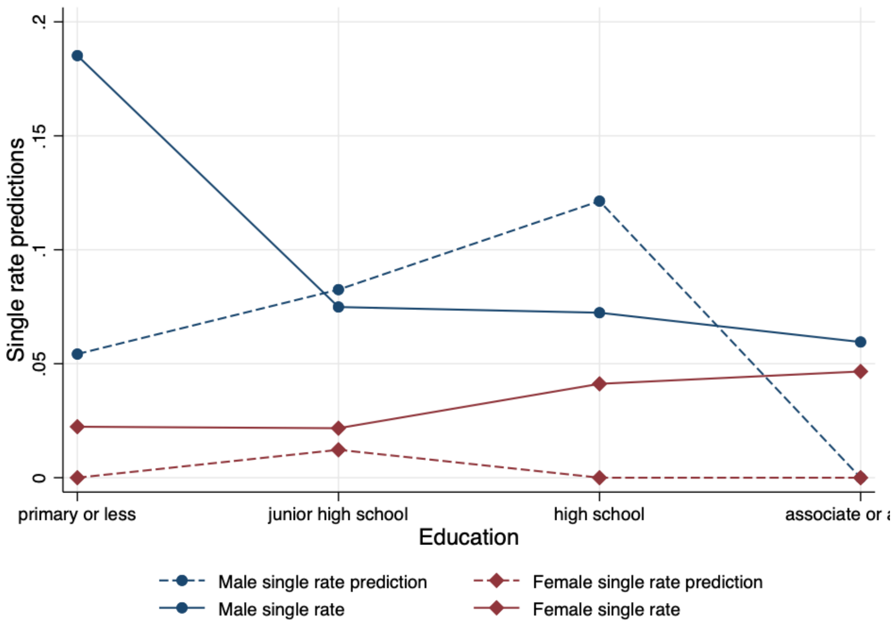
Note: Simulation 1 corresponds to the first counterfactual simulation where men's aversion to marrying a spouse more educated than themselves is turned off. Simulation 2 corresponds to the second counterfactual exercise where women's aversion to marrying a less educated spouse is turned off. For row "associate or above: primary or less", the entry stands for the number of females with associate or above education marrying males with primary or less education. Similar definitions apply to the rest of the table. Refer to Table 7, 8, 10, and 11 for full details on matching in the two counterfactual cases.

Figure 13: Single rate prediction: male's utility linear in spouse's education



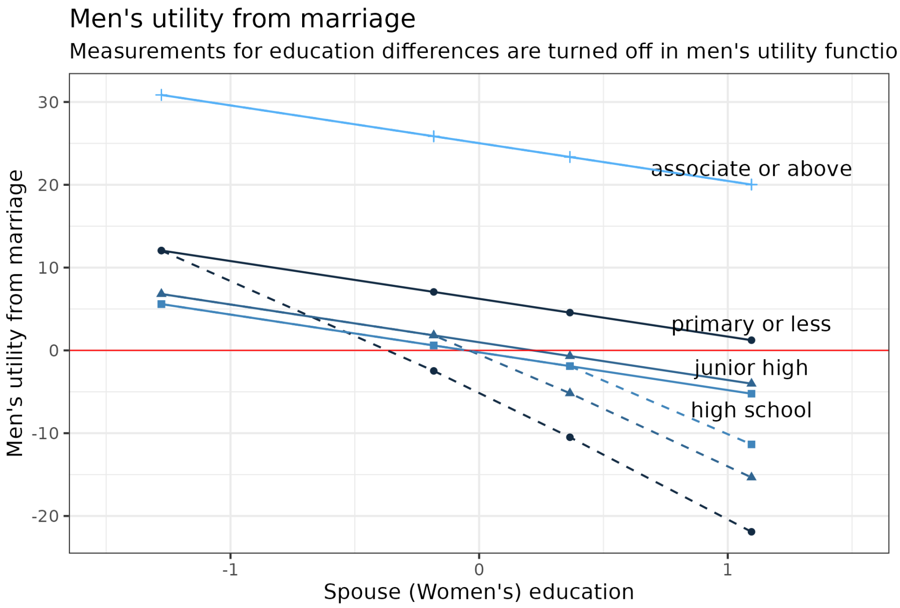
Note: This figure summarizes the prediction of single rates and the true single rates in the sample for the estimation.

Figure 14: Single rate prediction: female's utility linear in spouse's education

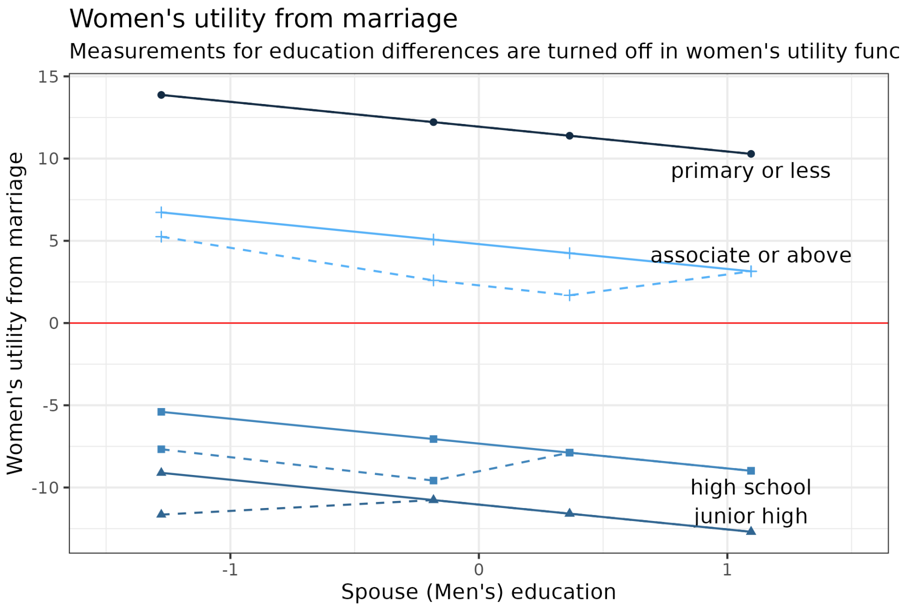


Note: This figure summarizes the prediction of single rates and the true single rates in the sample for the estimation.

Figure 15: Mean utilities from marriage [9]



(a) Counterfactual: men's utility from marriage



(b) Counterfactual: women's utility from marriage

Note: The solid lines represent the counterfactual mean utilities. The dashed lines are mean utilities from the baseline model.

[9] Refer to Table 9 in Appendix C.1 and Table 12 in Appendix C.2 for detailed numbers.

6 Are the rising single rates a pure result of changes in education distributions?

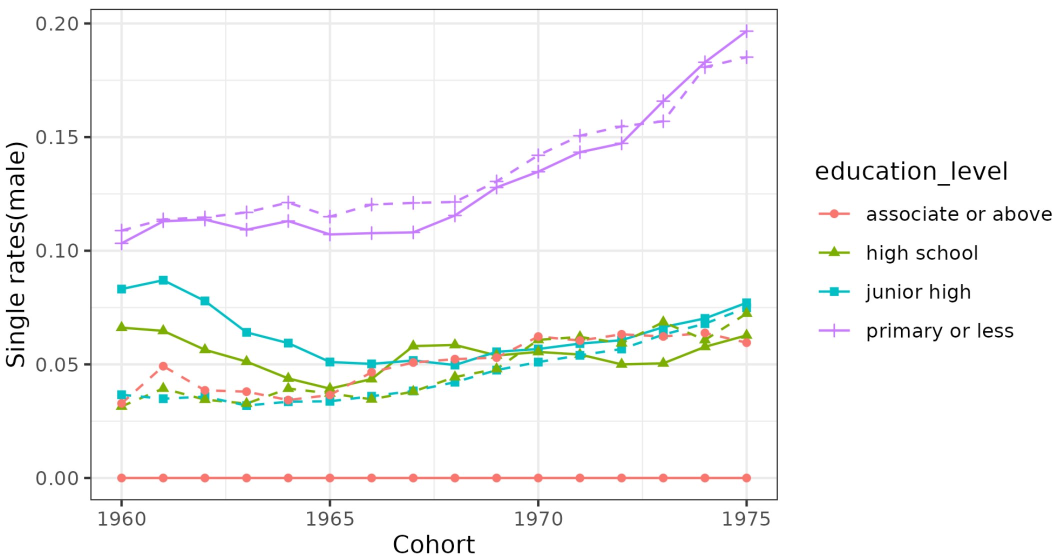
In this section, I ask whether the change in single rates across birth cohorts from 1960 to 1975 discussed in the beginning could be explained simply by the changes in education distributions, more specifically, whether the rising single rates of the least educated men and the most educated women are caused only by the changes in education distributions.

I construct the simulation sample for earlier cohorts in a manner similar to that used to estimate the baseline model parameters. I apply the estimated parameters from the baseline model, perform the matching following the deferred acceptance algorithm, and simulate the single rates.

In Figure 16 and 17, the dashed lines stand for true single rates from the Census, the solid lines are the predictions from the model. To smooth the simulated trend, the predicted single rates for each birth cohort are the average across the results from 30 random draws on taste parameters. The model pretty well predicts the rising single rates for men born between 1960 and 1975. It also fits the single rates for men with junior high school and high school education, especially for those born after 1965. For women, the fitting is much nosier. It predicts that women with associate education or more and high school education have higher singles than the other two less educated groups. However, high school-educated women are consistently predicted to have slightly higher single rates than the most educated group of women with associate or above education, which deviates from the patterns in the data. The result suggests that some important changes in tastes or matching technology played a significant role in the decrease in marriage among these women.

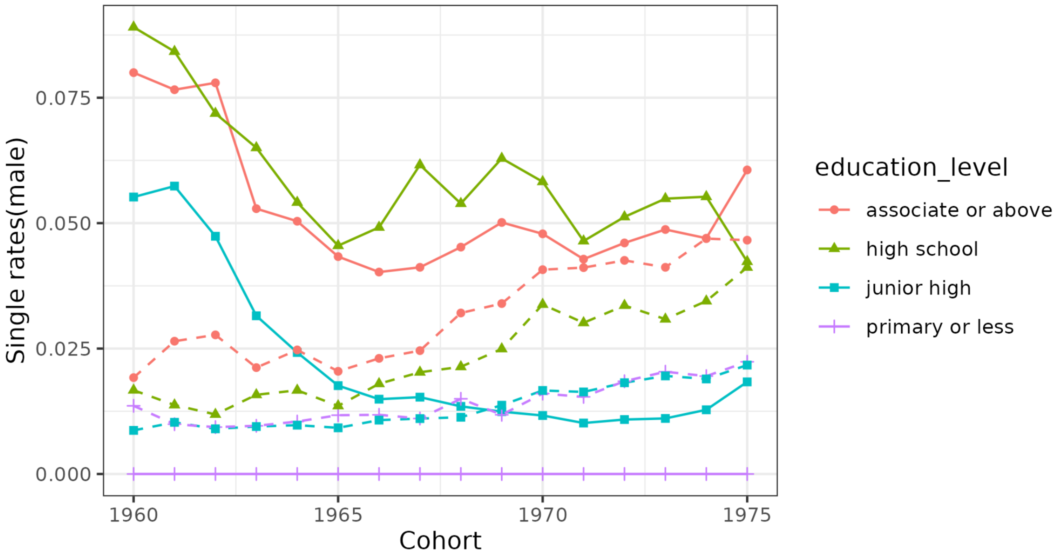
The results are robust even if I slightly change the number of cohorts used for simulation or draw random tastes a different number of times. The results from some other specifications for constructing the simulation samples (Figure 21, Figure 22, Figure 19 and Figure 20) are included in Appendix. Including more cohorts in one simulation helps smooth the trend, but the results do not differ much from patterns in the two figures here. The model's incapability of predicting any singles among the most educated men and least educated women persists for all birth cohorts and across all specifications.

Figure 16: Out-of-sample predicting: single rates for males



Note: This figure summarizes the predicted single rates for each earlier cohort. The dashed lines are true single rates based on Census data. The solid lines stands for the simulated single rates. To smooth the predicted trend, I mainly employ 4 cohorts, birth cohort which single rates are based on and the three birth cohorts before this cohort, when constructing the simulation sample of the married couples. This is the same as what I did in the baseline estimation. Each prediction is the average across the results from 30 random draws on taste parameters.

Figure 17: Out-of-sample predicting: single rates for females



Note: This figure summarizes the predicted single rates for each earlier cohort. The dashed lines are true single rates based on Census data. The solid lines stands for the simulated single rates. To smooth the predicted trend, I mainly employ 4 cohorts, birth cohort which single rates are based on and the three birth cohorts before this cohort, when constructing the simulation sample of the married couples. This is the same as what I did in the baseline estimation. Each prediction is the average across the results from 30 random draws on taste parameters.

7 Conclusion

In line with the literature that documented the husband's discounted utility when the wife earns more, like the work by Bertrand et al. (2015) and Chen and Hu (2021), I find evidence that suggests both men and women have aversions to marriages that consist of a more educated wife and a less educated husband for people born around 1972 to 1975 in China. This aversion is particularly pronounced among men, who usually assume the role of backbone for the household in the patriarchic social norm.

The distastes from both genders jointly shape the single rate patterns in the data. Missing aversions from either side, we will not see the unique pattern of men's single rate decreasing in education and women's increasing in education. Speaking to the high single rates of highly educated women that receive massive attention in the media, our analysis suggests two changes in preferences that could theoretically alleviate this issue: preventing the aversion to marrying more-educated women for men or preventing the aversion to marrying less-educated men for women. Seemingly symmetric, the implications on marriage matching are distinct, with single, well-educated women mostly marrying the least educated men in the first case and single, well-educated women mostly finding partners in the education levels right below them in the second scenario.

Along with women's empowerment in education, which is a change in the social landscape currently in motion in China and many other developing countries, as well as decades ago in developed countries, understanding the mismatch between marriageable men and women caused by the shifting education distributions and the enduring attitudes towards the gender role in a marriage becomes more relevant.

When more recent data is available, I can extend the analysis to see if the model could explain the marriage rates in these few years.

References

Almar, F., B. Friedrich, A. Reynoso, B. Schulz, and R. Vejlin (2023). Marital Sorting and Inequality: How Educational Categorization Matters.

Banerjee, A., E. Duflo, M. Ghatak, and J. Lafortune (2013). Marry for What? Caste and Mate Selection in Modern India. American Economic Journal: Microeconomics.

Bertrand, M., E. Kamenica, and J. Pan (2015, May). Gender Identity and Relative Income within Households*. The Quarterly Journal of Economics 130(2), 571–614.

Chen, Y. and D. Hu (2021, August). Gender norms and marriage satisfaction: Evidence from China. China Economic Review 68, 101627.

Gihleb, R. and K. Lang (2020, November). Educational Homogamy and Assortative Mating Have Not Increased *. In S. W. Polachek and K. Tatsiramos (Eds.), Research in Labor Economics, pp. 1–26. Emerald Publishing Limited.

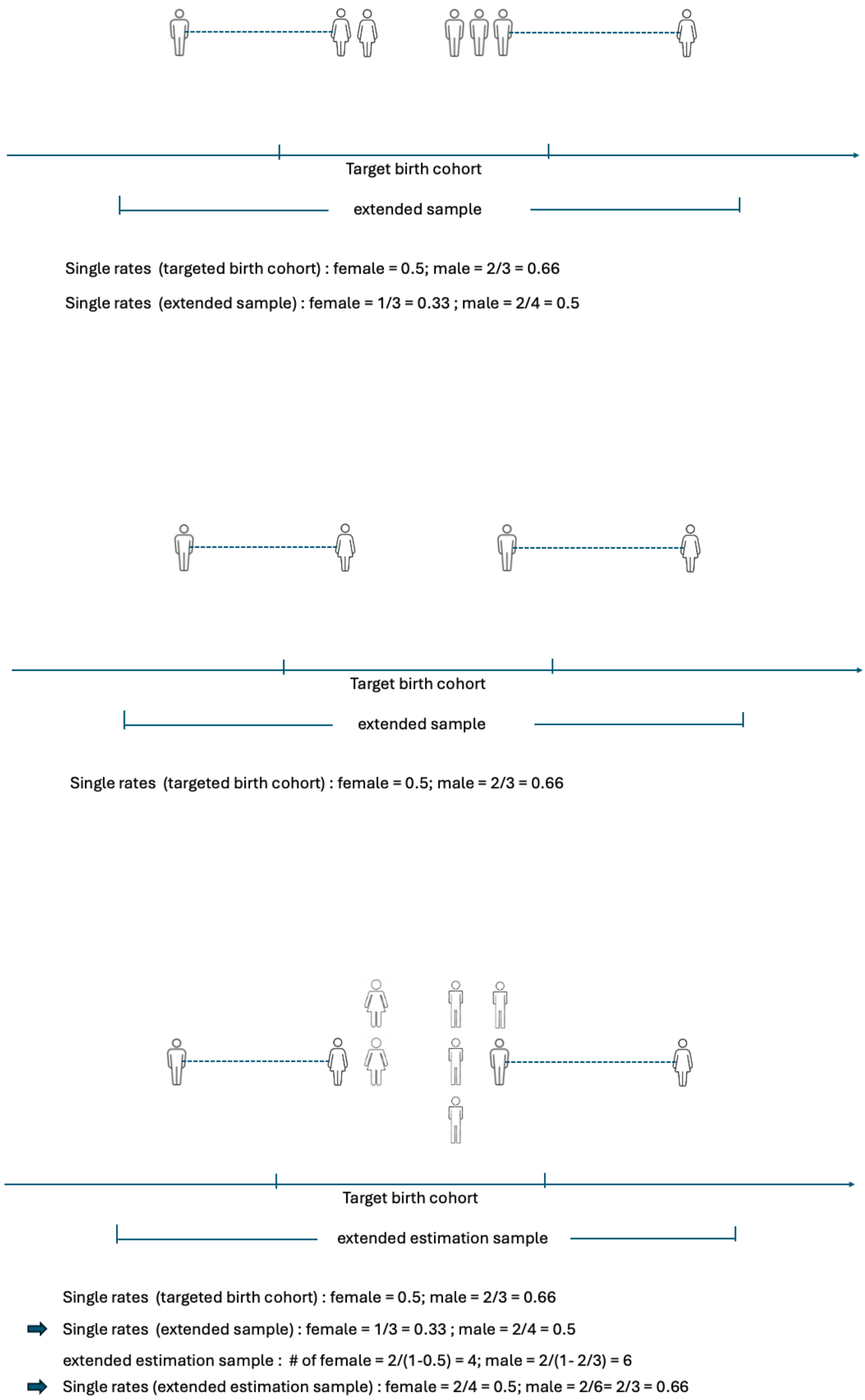
Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014). Marry Your Like: Assortative Mating and Income Inequality.

Han, L., T. Li, and Y. Zhao (2015, December). How Status Inheritance Rules Affect Marital Sorting: Theory and Evidence from Urban China. The Economic Journal 125(589), 1850–1887.

Appendix

A Sample construction

Figure 18: A simplified illustration of sample construction



B Baseline model

Table 6: Predicted mean utilities from marriage

Category	Male	Female
primary or less: primary or less	12.06	13.87
primary or less: junior high	-2.48	12.22
primary or less: high school	-10.49	11.39
primary or less: associate or above	-21.91	10.29
junior high: primary or less	6.81	-11.65
junior high: junior high	1.81	-10.76
junior high: high school	-5.18	-11.59
junior high: associate or above	-15.34	-12.69
high school: primary or less	5.59	-7.67
high school: junior high	0.60	-9.58
high school: high school	-1.90	-7.88
high school: associate or above	-11.35	-8.98
associate or above: primary or less	30.85	5.25
associate or above: junior high	25.86	2.59
associate or above: high school	23.36	1.69
associate or above: associate or above	20.03	3.15

Note: All the predicted utilities are rounded to the second decimal place. They are utilities from marriage relative to being single. For row "primary or less: junior high" and column "male", the entry stands for the utility a male with primary or less education can get from marrying a female with junior high school education, relative to the utility of a male with this education level being single. Similar definitions apply to the rest of the table.

C Counterfactual

C.1 Males' utility is linear in spouse's education

Table 7: Prediction precision: male matchings; male's utility linear in spouse's education

Category	Target	Prediction	Deviation
primary or less: primary or less	181.49	136.00	-45.49
primary or less: junior high	54.51	77.00	22.49
primary or less: high school	3.49	39.00	35.51
primary or less: associate or above	0.37	43.00	42.63
primary or less: single	54.51	0.00	-54.51
junior high: primary or less	174.84	196.00	21.16
junior high: junior high	630.03	541.00	-89.03
junior high: high school	40.32	60.00	19.68
junior high: associate or above	6.48	55.00	48.52
junior high: single	68.95	69.00	0.05
high school: primary or less	12.23	30.00	17.77
high school: junior high	93.75	113.00	19.25
high school: high school	93.56	43.00	-50.56
high school: associate or above	21.40	11.00	-10.40
high school: single	17.24	42.00	24.76
associate or above: primary or less	1.86	17.00	15.14
associate or above: junior high	20.14	81.00	60.86
associate or above: high school	43.25	47.00	3.75
associate or above: associate or above	128.19	56.00	-72.19
associate or above: single	12.24	5.00	-7.24

Note: All the numbers are rounded to the second decimal place. The estimation sample takes the sample with matched couples from the Census and scale down by 1/100 to keep time needed for matching in a reasonable range. This is the reason why the targets are not integers. Deviations less than 1 can be ignored and perfect predictions have been achieved. For row "primary or less: junior high", the entry stands for the number of females with primary or less education marrying males with junior high school education, as the target or as the predicted. Similar definitions apply to the rest of the table.

Table 8: Prediction precision: female matchings; male's utility linear in spouse's education

Category	Target	Prediction	Deviation
primary or less: primary or less	181.49	136.00	-45.49
primary or less: junior high	174.84	196.00	21.16
primary or less: high school	12.23	30.00	17.77
primary or less: associate or above	1.86	17.00	15.14
primary or less: single	8.47	0.00	-8.47
junior high: primary or less	54.51	77.00	22.49
junior high: junior high	630.03	541.00	-89.03
junior high: high school	93.75	113.00	19.25
junior high: associate or above	20.14	81.00	60.86
junior high: single	17.72	5.00	-12.72
high school: primary or less	3.49	39.00	35.51
high school: junior high	40.32	60.00	19.68
high school: high school	93.56	43.00	-50.56
high school: associate or above	43.25	47.00	3.75
high school: single	7.75	0.00	-7.75
associate or above: primary or less	0.37	43.00	42.63
associate or above: junior high	6.48	55.00	48.52
associate or above: high school	21.40	11.00	-10.40
associate or above: associate or above	128.19	56.00	-72.19
associate or above: single	7.64	0.00	-7.64

Note: All the numbers are rounded to the second decimal place. The estimation sample takes the sample with matched couples from the Census and scale down by 1/100 to keep time needed for matching in a reasonable range. This is the reason why the targets are not integers. Deviations less than 1 can be ignored and perfect predictions have been achieved. For row "primary or less: junior high", the entry stands for the number of females with primary or less education marrying males with junior high school education, as the target or as the predicted. Similar definitions apply to the rest of the table.

Table 9: Predicted mean utilities from marriage- male's utility linear in spouse's education

Category	Male	Female
primary or less: primary or less	12.06	13.87
primary or less: junior high	7.06	12.22
primary or less: high school	4.56	11.39
primary or less: associate or above	1.23	10.29
junior high: primary or less	6.81	-11.65
junior high: junior high	1.81	-10.76
junior high: high school	-0.69	-11.59
junior high: associate or above	-4.02	-12.69
high school: primary or less	5.59	-7.67
high school: junior high	0.60	-9.58
high school: high school	-1.90	-7.88
high school: associate or above	-5.23	-8.98
associate or above: primary or less	30.85	5.25
associate or above: junior high	25.86	2.59
associate or above: high school	23.36	1.69
associate or above: associate or above	20.03	3.15

Note: All the predicted utilities are rounded to the second decimal place. They are utilities from marriage relative to being single. For row "primary or less: junior high" and column "male", the entry stands for the utility a male with primary or less education can get from marrying a female with junior high school education, relative to the utility of a male with this education level being single. Similar definitions apply to the rest of the table.

C.2 Females' utility is linear in spouse's education

Table 10: Prediction precision: male matchings; female's utility linear in spouse's education

Category	Target	Prediction	Deviation
primary or less: primary or less	181.49	166.00	-15.49
primary or less: junior high	54.51	107.00	52.49
primary or less: high school	3.49	6.00	2.51
primary or less: associate or above	0.37	0.00	-0.37
primary or less: single	54.51	16.00	-38.51
junior high: primary or less	174.84	182.00	7.16
junior high: junior high	630.03	562.00	-68.03
junior high: high school	40.32	95.00	54.68
junior high: associate or above	6.48	6.00	-0.48
junior high: single	68.95	76.00	7.05
high school: primary or less	12.23	23.00	10.77
high school: junior high	93.75	103.00	9.25
high school: high school	93.56	66.00	-27.56
high school: associate or above	21.40	18.00	-3.40
high school: single	17.24	29.00	11.76
associate or above: primary or less	1.86	8.00	6.14
associate or above: junior high	20.14	35.00	14.86
associate or above: high school	43.25	22.00	-21.25
associate or above: associate or above	128.19	141.00	12.81
associate or above: single	12.24	0.00	-12.24

Note: All the numbers are rounded to the second decimal place. The estimation sample takes the sample with matched couples from the Census and scale down by 1/100 to keep time needed for matching in a reasonable range. This is the reason why the targets are not integers. Deviations less than 1 can be ignored and perfect predictions have been achieved. For row "primary or less: junior high", the entry stands for the number of females with primary or less education marrying males with junior high school education, as the target or as the predicted. Similar definitions apply to the rest of the table.

Table 11: Prediction precision: female matchings; female's utility linear in spouse's education

Category	Target	Prediction	Deviation
primary or less: primary or less	181.49	166.00	-15.49
primary or less: junior high	174.84	182.00	7.16
primary or less: high school	12.23	23.00	10.77
primary or less: associate or above	1.86	8.00	6.14
primary or less: single	8.47	0.00	-8.47
junior high: primary or less	54.51	107.00	52.49
junior high: junior high	630.03	562.00	-68.03
junior high: high school	93.75	103.00	9.25
junior high: associate or above	20.14	35.00	14.86
junior high: single	17.72	10.00	-7.72
high school: primary or less	3.49	6.00	2.51
high school: junior high	40.32	95.00	54.68
high school: high school	93.56	66.00	-27.56
high school: associate or above	43.25	22.00	-21.25
high school: single	7.75	0.00	-7.75
associate or above: primary or less	0.37	0.00	-0.37
associate or above: junior high	6.48	6.00	-0.48
associate or above: high school	21.40	18.00	-3.40
associate or above: associate or above	128.19	141.00	12.81
associate or above: single	7.64	0.00	-7.64

Note: All the numbers are rounded to the second decimal place. The estimation sample takes the sample with matched couples from the Census and scale down by 1/100 to keep time needed for matching in a reasonable range. This is the reason why the targets are not integers. Deviations less than 1 can be ignored and perfect predictions have been achieved. For row "primary or less: junior high", the entry stands for the number of females with primary or less education marrying males with junior high school education, as the target or as the predicted. Similar definitions apply to the rest of the table.

Table 12: Predicted mean utilities from marriage- female's utility linear in spouse's education

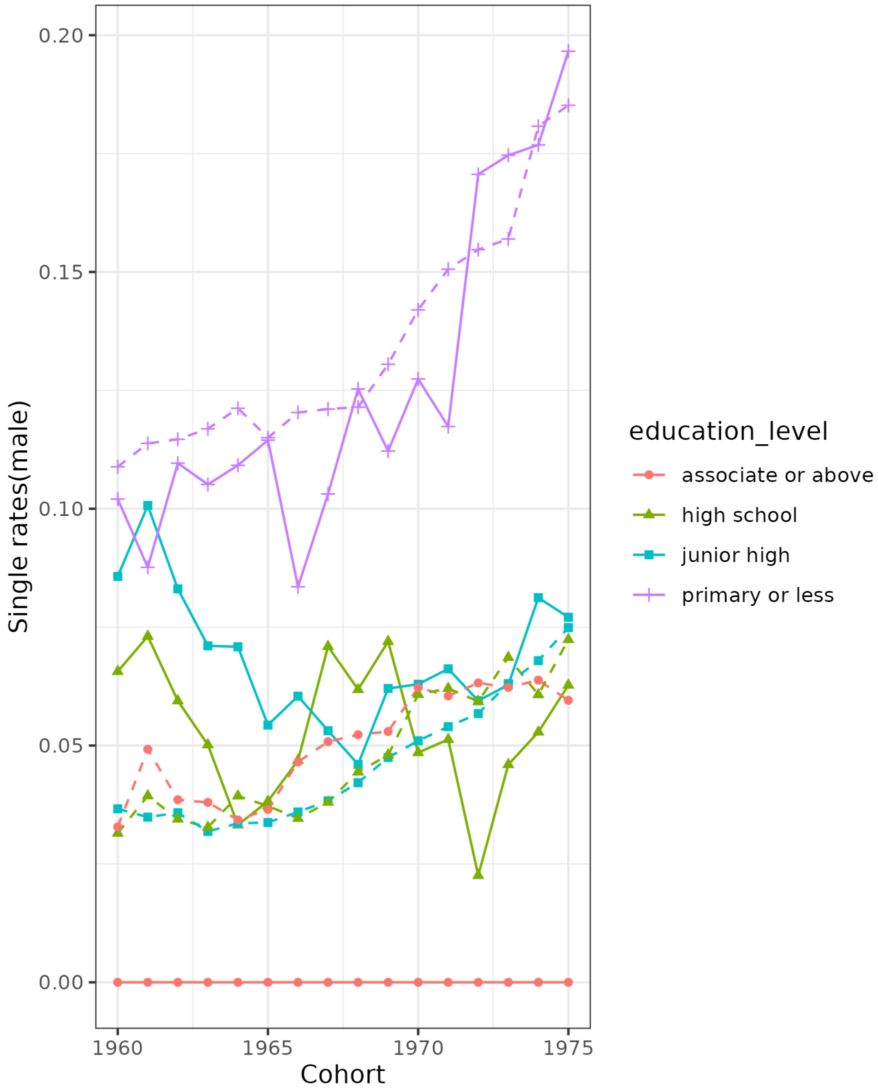
Category	Male	Female
primary or less: primary or less	12.06	13.87
primary or less: junior high	-2.48	12.22
primary or less: high school	-10.49	11.39
primary or less: associate or above	-21.91	10.29
junior high: primary or less	6.81	-9.11
junior high: junior high	1.81	-10.76
junior high: high school	-5.18	-11.59
junior high: associate or above	-15.34	-12.69
high school: primary or less	5.59	-5.40
high school: junior high	0.60	-7.05
high school: high school	-1.90	-7.88
high school: associate or above	-11.35	-8.98
associate or above: primary or less	30.85	6.73
associate or above: junior high	25.86	5.08
associate or above: high school	23.36	4.25
associate or above: associate or above	20.03	3.15

Note: All the predicted utilities are rounded to the second decimal place. They are utilities from marriage relative to being single. For row "primary or less: junior high" and column "male", the entry stands for the utility a male with primary or less education can get from marrying a female with junior high school education, relative to the utility of a male with this education level being single. Similar definitions apply to the rest of the table.

D Different specifications for out-of-sample prediction

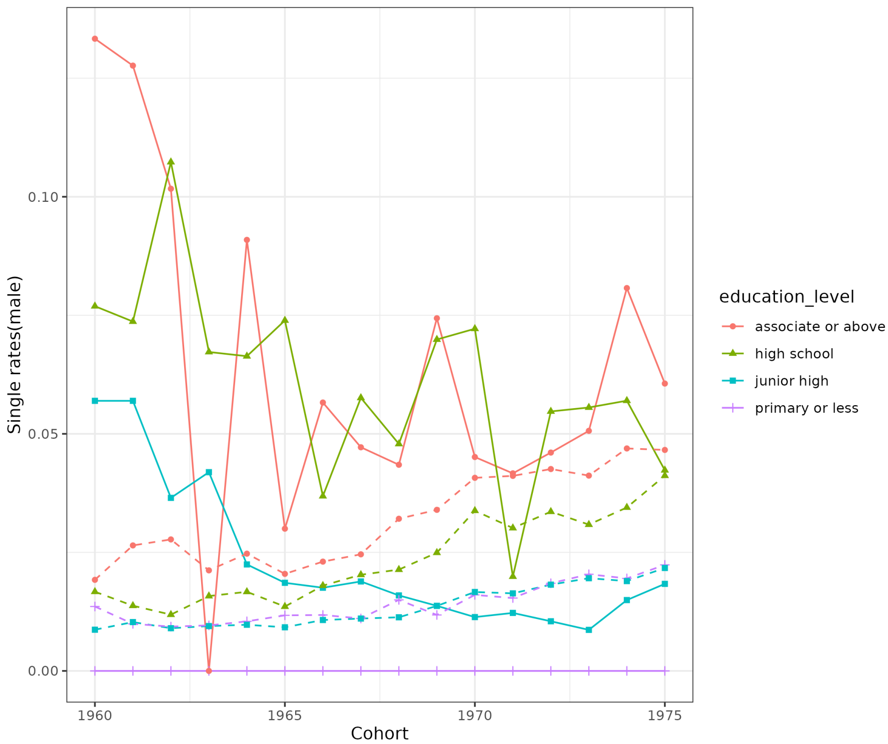
D.1 sample construction: [target cohort -3, target cohort]; sets of random taste parameters: 1

Figure 19: Out-of-sample predicting: single rates for males



Note: This figure summarizes the predicted single rates for each earlier cohort, using the estimated preferences and following the procedure of constructing estimation sample and matching in the baseline model .

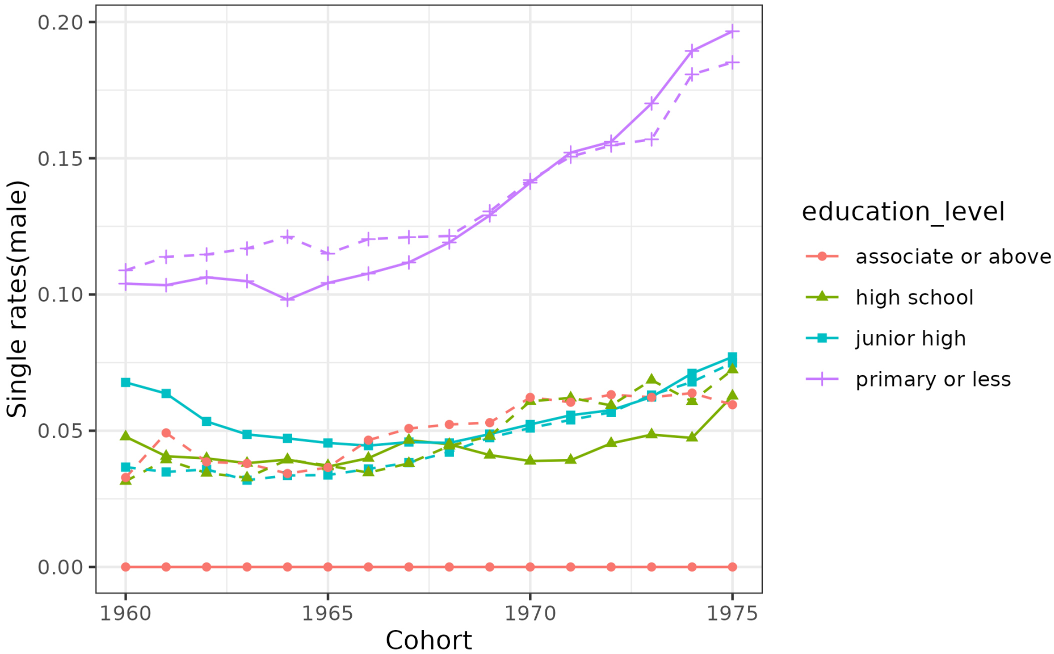
Figure 20: Out-of-sample predicting: single rates for females



Note: This figure summarizes the predicted single rates for each earlier cohort, using the estimated preferences and following the procedure of constructing estimation sample and matching in the baseline model.

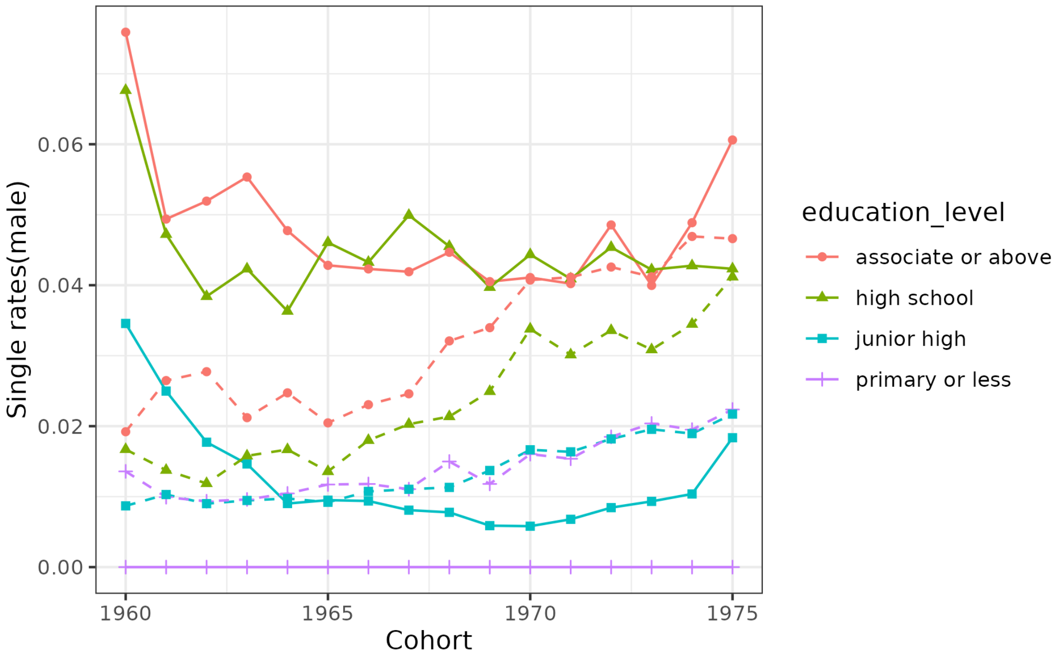
D.2 sample construction: [target cohort -2, target cohort +2]; sets of random taste parameters: 30

Figure 21: Out-of-sample predicting: single rates for males



Note: This figure summarizes the predicted single rates for each earlier cohort. The dashed lines are true single rates based on Census data. The solid lines stands for the simulated single rates. To smooth the predicted trend, I mainly employ 5 cohorts when constructing the simulation sample of the married couples, with 2 cohorts before and 2 cohorts after the respective year for which single rates are calculated. Each prediction is the average across the results from 30 random draws on taste parameters.

Figure 22: Out-of-sample predicting: single rates for females



Note: This figure summarizes the predicted single rates for each earlier cohort. The dashed lines are true single rates based on Census data. The solid lines stands for the simulated single rates. To smooth the predicted trend, I mainly employ 5 cohorts when constructing the simulation sample of the married couples, with 2 cohorts before and 2 cohorts after the respective year for which single rates are calculated. Each prediction is the average across the results from 30 random draws on taste parameters.