Social Status and Marriage Markets: Evaluating a *Hukou* Policy in China*

Qinyou Hu†

Click here for the latest version

Abstract

Social status is a key determinant of marital sorting patterns. This paper studies the role of the *hukou* system — a household registration system in mainland China that categorizes individuals into rural and urban groups — in shaping marital sorting patterns. Specifically, I evaluate a policy that granted men the same rights to transfer their *hukou* status to children as women. The policy significantly increased *hukou* intermarriages, particularly between rural women and urban men. By estimating a two-sided directed search and matching model, I find that the policy distorted the marriage market by increasing search frictions to a greater extent for rural men than for urban men. Highly educated rural men would have benefited the most if this policy had not been implemented, indicating that they suffered the most welfare loss, at least in terms of marriage outcomes. The findings underscore the importance of understanding the role of social status in shaping marriage markets and the impact of status-related policy interventions on welfare outcomes.

Keywords: Social status, Marriage market, Empirical matching, Gender
JEL classification: D1, J12, J18

*I am grateful to Peter Arcidiacono, Andrew Beauchamp, Rossella Calvi, Flávio Cunha, Chao Fu, Naijia Guo, Sakshi Gupta, Yinghua He, Ruixue Jia, Bolun Li, Elaine Liu, David Ong, Isabelle Perrigne, Andrea Salvati, Xun Tang, Fisher Yu, Danyan Zha, Junsen Zhang, Shumeng Zhang, and participants at the HCEO Summer School on Socioeconomic Inequality in Chengdu, SOLE 2021 Conference, AASLE 2021 Conference, Young Economist Symposium 2021, 2021 Symposium on Contemporary Labor Economics at Xiamen University, and Rice Economics Department Seminar for valuable comments and suggestions. All errors are my own.

†Department of Economics, Rice University. Email: qinyou.hu@rice.edu.

1 Introduction

In many societies, individuals' social status influences their educational and professional opportunities, their interpersonal relationships, and even who they choose as a spouse or life partner. One way that social status is determined is through rural vs. urban regional distinctions, a particularly prominent feature in the developing world. Recent literature has shown that about half the inequality in the developing world can be attributed to the rural-urban gap (Lagakos, 2020). China is a special case in that the rural-urban divide is institutionalized by the *hukou* system, a unique household registration system that labels people based on their birth locations. Because the localized public welfare system is closely tied with the *hukou* system, urban *hukou* confers more economic benefits than rural *hukou*, including free access to better-quality public schools, a higher possibility of securing a prestigious occupation, and more advanced healthcare facilities (Song, 2014; Wu and Treiman, 2004). Due to this linked value difference, the institutional dichotomy has widened the disparity and inequality between rural and urban populations in China ever since the 1950s.[1] Studies have shown that the current *hukou* system is also exacerbating inequalities between different *hukou* holders in the marriage market (Meng and Zhao, 2019; Qian and Qian, 2017; Sharygin et al., 2013; Yu and Xie, 2015).

My paper adds to the previous literature by examining the effect of a unique status inheritance rule, which significantly changed in 1998, on marital sorting in China. The *hukou* status, which is hereditary and challenging to alter later in life, underwent an important legislative alteration in 1998: State Council Order 1998 No. 24 granted men the same rights as women in passing on their *hukou* status to their children. Before 1998, children could only inherit their mother's *hukou*, but now they can inherit either parent's *hukou*. Given that children are a crucial outcome of marriage, status inheritance rules are likely to influence marital decisions. Analogous to the caste system in India, the *hukou* system represents a critical aspect of social identity in China (Afridi et al., 2015). As such, trends in *hukou* intermarriages, i.e., marriages between urban and rural *hukou*, serve as a vital indicator of social mobility patterns and have significant implications for the issue of inequality.

Scholarly literature on the effects of this status inheritance rule change on marriage markets is scarce. This paper specifically extends the current literature by answering the following questions. i) What role does social status play in the marriage market? How does the 1998 status inheritance rule policy affect marriage decisions, especially *hukou* intermarriages? (ii) Do men and women value social status differently in the marriage matching process? Is it possible to separately identify the effects of the status inheritance rule change policy on men and women? iii) Given the model estimates, how would the marriage market change if the

[1] This phenomenon has drawn attention from scholars in various academic fields, including economics (e.g., Afridi et al., 2015; Li et al., 2017; Meng, 2012; Tombe and Zhu, 2019), sociology (e.g., Bian, 2002; Wang and Schwartz, 2018; Wu, 2019; Wu and Treiman, 2004), political science (e.g., Vortherms, 2019, 2021), and public health (e.g., Gong et al., 2012).

1998 policy had not been implemented? What would happen if there were no such differences in *hukou* status?

To answer the first research question, I conduct a difference-in-differences (DID) analysis using administrative *census* data to explore patterns in the policy's impacts on *hukou* intermarriages. This exercise seeks to generate suggestive evidence to motivate the structural model. I construct the first difference based on whether the groom was younger than the legal marriage age, 22, in 1998, and the second difference measures the population gap between urban and rural *hukou* holders in each province before the reform. This second layer of difference can capture the scope of the policy's impact on intermarriages. Intuitively, if there is a significantly smaller urban population than rural population within a province, or vice versa, there would be a smaller likelihood of intermarriages.

The estimates show the reform led to a 5-percentage-point, statistically significant increase in intermarriages between urban and rural *hukou* holders, with the rate of intermarriage growing slower for urban women than for urban men. This result is consistent with previous findings from Han et al. (2015) and Nie and Xing (2011). Given the initial intermarriage rate of around 6.5%, a 5-percentage-point increase is equivalent to a substantial rise of 76.9% in the number of intermarriages observed within the sample of married individuals.

Though these findings confirm that the status-related policy does influence marriage decisions, it is unclear why men and women behave differently. The data exploration exercise following the DID analysis is also limited in identifying individuals' preferences for *hukou* status and in capturing the general equilibrium effect of market conditions, such as the corresponding sex ratios or potential frictions, on marriage outcomes. Furthermore, this exercise is unable to conduct welfare analysis for policy evaluation.

To address these analytical limitations, I further develop and estimate a two-sided, directed search and matching model to answer the second research question: Do men and women value *hukou* status differently in the marriage matching process? The model aids in uncovering the underlying mechanisms of the policy's impact, and can be used to analyze the general equilibrium effect of potential changes in marriage market conditions due to the 1998 status inheritance rule change. I incorporate the search process for two reasons. First, it takes time to find a desirable match. Hindered by the segregation generated by the *hukou* system (Meng and Zhang, 2001), social contact between urban and rural *hukou* individuals is limited, especially for rural individuals looking for an urban spouse. Second, the search process can accurately reflect the realities of the Chinese marriage market. Family members, usually the parents, frequently recommend and even decide on marriage partners without the consent of the younger individuals involved, potentially causing information friction.[2] It is a directed search as marital

[2]The emergence of matchmaking corners ("xiangqin jiao") in contemporary China, often meeting in parks and arranged by parents (Wang, 2021), suggests that parental participation still influences young couples' marriage choices to some extent. Information friction could happen because families might provide biased information on potential spouses to their adult chil-

searches tend to be targeted based on observed characteristics (e.g., Cornelius, 2003; Ong et al., 2020; Shimer and Smith, 2000; Wright et al., 2021).

My model allows individuals to match along multidimensional characteristics, including educational attainment and *hukou* status. Following Arcidiacono et al. (2016), I embed the model within a non-transferable utility (NTU) framework to disentangle the heterogeneous impacts of the 1998 status inheritance rule change on men and women, rather than using the transferable utility (TU) framework from the matching literature, which can only identify the joint marital surplus or gains from marriage. The equilibrium derived from the model is characterized by a fixed point solution to the searching and matching probability functions. The intuition behind it involves the interplay of the matching and searching probabilities. People make a search decision based on the probability of matching success. The number of searches in the model would, in turn, affect the competition level in the marriage market and thus the matching probability.[3] Unlike Arcidiacono et al. (2016), my model incorporates a more comprehensive specification of the matching technology function to capture the impact of the policy, particularly the diverse patterns of the two intermarriage types motivated by the DID analysis. Specifically, I allow the 1998 policy to play a role in the search friction in the marriage matching technology function as well.

I estimate the model using administrative data from the 2005 *Mini-Census* and aggregate measures from the 2000 *Census* published by the National Bureau of Statistics of China. The data pattern shows that most marriages occur between spouses with similar characteristics, which supports the phenomenon of positive assortative mating proposed in Becker (1973). In terms of preferences for education, the model estimates suggest that both men and women prefer high educational attainment. Interestingly, I find that women prefer higher educational attainment in their spouses, whereas men dislike it. After accounting for the heterogeneity across *hukou* types, urban women have a stronger preference than rural women for having the same educational attainment as their spouses. On the other hand, rural men tend to hold this preference more strongly than urban men. When it comes to preferences for *hukou* status, urban men favor women with rural *hukou*, but urban women dislike men with rural *hukou*, suggesting that urban women may value a higher social status more than urban men in marriage decisions, especially when women are in short supply in the market.[4] I also find that women place a higher value on sharing similar characteristics with their partners than men do. Women, in general, prefer men with urban *hukou* status. With respect to matching technology parameters,

dren, or might make sub-optimal marriage decisions on behalf of their children (Raiber et al., 2023). For more details on these matchmaking corners in China, see Vandenberg (2018).

[3]Thanks to an anonymous referee for pointing this out. One can interpret the number of searches in the model as the number of relationships or the number of attempts made to ask someone out in real-life scenarios, which can be further verified by more detailed data.

[4]Similarly, in another context, Abramitzky et al. (2011) found that greater male scarcity in France after WWI led to improved marital prospects for men, as they were less likely to marry women of lower social classes.

I find that rural men, relative to urban men, indeed experienced significantly higher friction in marriage markets after the policy.

The added value of the model is that it enables counterfactual policy simulations to calculate the corresponding changes in welfare. I thus proceed to answer the third research question, how the marriage market would have changed under a range of scenarios, using the model estimates. To better understand the policy's impacts on welfare changes in marriage outcomes, I simulate a scenario in which the 1998 policy had not been implemented, the study's main counterfactual policy experiment. By removing the policy indicator in the matching technology, I am able to predict the marriage matching pattern, particularly the *hukou* intermarriage rate, using the structural model. Additionally, I perform simulation exercises under three other counterfactual scenarios. First, in order to better understand the trade-off between educational attainment and social status, I increase all rural women's educational attainment by equipping them with the same distribution of educational attainment as urban women. The same exercise is carried out on men for comparison, to depict the gender disparity. Second, in line with the recent fertility relaxation policy in China near the end of 2015,[5] which may have resulted in changes to the sex ratio, I investigate the impact of changing the sex ratio on marriage matching by increasing and decreasing the number of females in each province by 10 percent. I also use the WHO's formal definition of the natural sex ratio, which is 105 males for every 100 females, to simulate a marriage market with a balanced sex ratio.[6] Third, I completely eliminate the *hukou* status differences in order to mimic the purpose of the 2014 national *hukou* reform – to gradually remove the *hukou* system.[7] As a result, individuals will only differ in dimensions other than their *hukou* status when entering the marriage market. I calculate the welfare change for all types of individuals following each of the policy experiments. As my model does not encompass other dimensions of change, such as changes in labor markets or migration patterns, the welfare metrics based on the model estimates are relatively conservative.

In the absence of the 1998 policy change, the structural model predicts a decrease in intermarriages but more marriages between rural men and rural women. The welfare calculation implies that rural and highly educated men would benefit the most, while urban and less-educated men would lose the most.[8] Given the various proportions of each type of individual, I compute

[5] China abandoned its long-lasting *One Child Policy* at the end of 2015, and the "universal two-child policy" took effect in January 2016. China Daily (2017). "Most births this year added a child beyond the first one." http://www.chinadaily.com.cn/china/2017-07/24/content_30219626.htm.

[6] Although the most relevant sex ratio for the analysis is that of adults, it remains challenging to establish a consensus on this variable due to age-related changes in mortality rates among men and women (e.g., Wang et al., 2012; Zhu et al., 2019). As a result, I use the official statistics from the WHO, which reflect the sex ratio at birth, available at https://ourworldindata.org/gender-ratio.

[7] The official document (in Chinese) can be accessed here: The CPC Central Committee and the State Council, *Guojia xinxing chengshihua guihua (2014-2020 nian)* (National Plan on New Urbanization [2014-2020]), issued on 16 March 2014, http://www.gov.cn/zhengce/2014-03/16/content_2640075.htm.

[8] In my analysis, I define high education as having completed high school or above, and low education as having completed primary school or less. See more details in Section 2.2.

the overall welfare change in the absence of the policy reform. The result is a net increase, indicating that the 1998 *hukou* inheritance amendment ended up exacerbating inequality and eroding the nation's overall welfare, at least from the perspective of marriage markets. Without the differences in *hukou* status, I find that provinces with higher urban populations have a more substantial proportion of "losers," i.e., those who would experience welfare loss in the counterfactual simulation but not in the baseline condition. Women would benefit more than men from removing the *hukou* differences. After analyzing the welfare changes for each counterfactual exercise, I find that the policy that completely eliminates the *hukou* differences is the only one that would generate welfare gains for all individuals in the marriage market. These findings reveal a significant welfare distortion in the Chinese marriage market, driven by the existing notoriously imbalanced sex ratio and the substantial disparity established by the *hukou* system, which cannot be captured solely by the DID analysis.

Related Literature. This study is among the first to examine the effect of urban status-related policies on marriage markets through the lens of a directed search and matching model. Only three studies have investigated this particular status inheritance rule change policy to date – Nie and Xing (2011), Han et al. (2015), and Han and Shi (2019). Both Nie and Xing (2011) and Han et al. (2015) studied the policy impacts on the marriage matching pattern, while Han and Shi (2019) examined how the policy affects the intrahousehold dynamics of married couples and the welfare of their offspring. Like Nie and Xing (2011) and Han et al. (2015), my paper also finds a significant increase in inter-*hukou* marriages as a result of the policy change. However, my study is able to identify marriage matching preferences for both genders, rather than just depicting changes in marital matching outcomes. Understanding these heterogeneous preference patterns for men and women is crucial for comprehending how the abolition of the matrilineal rule contributes to gender inequality in Chinese society. Furthermore, my model accounts for search friction, a crucial feature of marriage markets in the developing world, a key analytical feature that is absent from the previous three studies. By means of structural model estimation, this study quantifies the welfare consequences of status-related policies from the perspective of the marriage market. The findings offer valuable insights for the current *hukou* relaxation reform in China, as well as for other developing countries where social status may be a prominent consideration.

My paper closely relates to the literature on how social status influences marriage decisions (Abramson and Shayo, 2022; Afridi et al., 2015; Akerlof and Kranton, 2000; Atkin et al., 2021; Fershtman et al., 1996; Oh, 2021; Weiss and Fershtman, 1998). The prevailing view is that social status contributes to understanding inequality and is non-negligible in determining marriage outcomes (Adda et al., 2020; Banerjee et al., 2013; Kalmijn, 1991). This study contributes to the literature by examining the unique *hukou* system in China. Like Afridi et al. (2015), my paper also considers *hukou* as a social identity. In contrast to caste or race, *hukou*

status includes both economic and non-economic aspects, potentially generating novel insights about marriage markets. Previous literature on the role of *hukou* in marriage decisions has found that the *hukou* system creates inequality in marriage outcomes, as reflected in people's preferences.[9] This study expands that observation by evaluating the impact of a unique status inheritance policy change on urban and rural *hukou* intermarriages. Through a matching model, this paper considers other crucial aspects of the Chinese marriage market, such as the imbalanced sex ratio and market frictions, thus providing a comprehensive overview of the policy's welfare impact.

My work also adds to the growing literature on marriage matching models. Most marriage matching studies applied the frictionless matching framework (Adda et al., 2020; Brandt et al., 2018; Chiappori, 2017; Chiappori et al., 2017; Chiappori and Salanié, 2016; Dupuy, 2021; Dupuy and Galichon, 2014; Han et al., 2015).[10] This paper differs by incorporating search friction, a unique and critical feature of the Chinese marriage market, to study intermarriages between holders of different *hukou* status. Friction is prevalent and unavoidable, as described in Chiappori and Salanié (2021), and must be considered when studying marriage decisions in developing countries. Although there has been substantial theoretical work on directed search methods(e.g., Chade et al., 2017; Wright et al., 2021), very little theoretical and empirical literature applies them in the marriage market.[11] Instead, I use observational data and impose functional form assumptions on the model structure to study search and matching behavior in marriage decisions. Additionally, the unique status inheritance rule change serves as an exogenous variation to help with model identification and estimation. There are a few studies that combine TU and search framework to study marital matching (e.g., Shimer and Smith, 2000). With the NTU framework, this study goes a step further by demonstrating gender heterogeneity in matching preferences and allowing for trade-offs of multiple traits. The gendered disparity in preferences plays a role in explaining marriage dissolution and has further implications for intra-household inequality (Falk and Hermle, 2018).

Third, my research contributes to the literature studying the impact of the sex ratio on marriage outcomes. Seminal work by Becker (1973, 1974, 1981) was the first to incorporate

[9]For example, sociological literature linked *hukou* intermarriages to socioeconomic mobility in China (e.g., Wang and Schwartz, 2018; Yu and Xie, 2015), and explored the impact of rural-urban migration on *hukou* intermarriages (e.g., Dupuy, 2021; Meng and Zhao, 2019).

[10]Choo and Siow (2006), for example, empirically estimated a static TU matching model and embedded marriage matching into a discrete choice framework. Chiappori et al. (2012) extended the framework of Choo and Siow (2006) and built a marriage matching model on multiple dimensions by summarizing them into a one-dimensional index. Brandt et al. (2018) employed both static and dynamic versions of the marriage matching framework proposed by Choo and Siow (2006) and Choo (2015) to study the "leftover women" phenomenon in China, and estimated the gains to marriage for two-dimensional matching on age and education. Their findings suggest that the term "leftover women" may not be an accurate description, as these women are not truly left behind but are rather marrying later. Dupuy (2021) applied a TU matching model and a hedonic model to quantify the relative importance of labor and marital motives for migration in China.

[11]One challenge is data limitations, as it is usually challenging to detect search behaviors in marriage decisions. To tackle this problem, a few studies use lab or field experiments to track people's search and matching behaviors (Hitsch et al., 2010a,b; Ong and Wang, 2015; Ong et al., 2020).

the notion of the family into economics studies. Becker argued that marriage decisions can be modeled as multi-agents' actions in segmented "markets" defined by various personal attributes, such as age, educational attainment, etc., and that men and women can be considered as the supply and demand sides of these "markets." As such, changes in the sex ratio play a crucial role in determining marriage matching outcomes. Later research has devoted considerable attention to understanding marriage markets.[12] The present study also empirically tests the impact of sex ratios on marriage matching patterns. In addition, it incorporates the role of marriage market competitiveness (skewed sex ratios) as well as the search costs in evaluating the impact of status-related policies on the Chinese marriage market. Prior research has linked the skewed sex ratio in China with the son preference tradition and the well-known *One Child Policy* (Ebenstein, 2011) and discussed the further implications for marriage markets (Edlund, 1999; Guilmoto, 2011; Porter, 2016).[13] This study enriches the literature by capturing the effects of both the sex ratio and *hukou* status on the marriage market at the same time. The counterfactual exercise also helps understand how the relaxation of the *One Child Policy*, combined with the abolition of *hukou* status, would reshape the Chinese marriage market.

Paper Structure. The remainder of the paper is organized as follows. In Section 2, I introduce the institutional background and unique features of the Chinese marriage market. I also detail the data sources and present the motivating evidence. In Section 3, I describe the model setup, the identification, and the estimation of the marriage matching model. In Section 4, I report the estimation results and conduct counterfactual exercises. Section 5 concludes.

2 Data and Empirical Evidence

2.1 Background

2.1.1 The Chinese *Hukou* System

In 1958 China implemented the *hukou* system, a unique household registration system. The system was created with the goal of regulating migration and allocating resources efficiently for the entire country (Meng, 2012; Xiang, 2015). Two dimensions of constraints are imposed by the *hukou* system: type and location. The type refers to "agricultural" and "non-agricultural," also known as rural and urban *hukou*, respectively (Song, 2014). The location is strongly linked to a person's presumed permanent or regular residence, such as a city, town, or village. *Hukou* also has a hereditary feature: when children are born, they automatically inherit their parents' *hukou* type and location.

[12] On the economics of marriages, see the reviews by Weiss (1997) and, more recently, Browning et al. (2014).

[13] Edlund (1999), for example, developed a theoretical model to analyze the impact of prenatal sex determination due to the son preference tradition on later marital matching patterns, and predicted that a skewed sex ratio can lead to women marrying socially superior men.

As the ultimate record of identity, one's *hukou* is the prerequisite document necessary to obtain any other form of identity, other than a birth certificate. Moreover, the possession of a *hukou* entitles citizens to government services (Vortherms, 2019). More crucially, *hukou* status is directly linked to the public welfare system. Different types are tied to different public services, welfare benefits, and educational opportunities. For instance, the majority of elite schools and high-quality hospitals in China are located in large cities, so having a rural *hukou* in a small city is usually associated with low-quality benefits compared to having one in a big city.

Unlike most Western countries, where rural-to-urban migration is unrestricted, the Chinese government institutionalizes the urban-rural divide and exerts stringent control over urban-rural boundaries. It is exceedingly difficult to convert from rural to urban *hukou* (Song, 2014; Wu and Treiman, 2004; Xiang, 2015). Since the early 1980s, China has gradually enacted a series of laws to delegate fiscal and administrative powers, including the management of the *hukou* systems, to lower-level administrations. The requirements for *hukou* conversion are set by city governments and vary at the municipal level. The overall trend is for cities to grant local status to those who are either wealthy, i.e., able to purchase urban housing or make substantial investments, or who have relatively high skills, such as a college diploma (at the minimum) or a professional qualification. To fulfill the requirement of *hukou* conversion, one must satisfy multiple criteria at the same time.

2.1.2 The 1998 *Hukou* Inheritance Law Reform

One of the most distinguishing features of *hukou* is that it is hereditary, meaning that children's *hukou* is passed down directly from their parents. China enacted a nationwide *hukou* inheritance law change in 1998.[14] Before August 1998, a child's *hukou* status had to follow his or her mother, whereas children born after 1998 were allowed to freely adopt either their mother's or father's *hukou* identity (State Council Order 1998 No. 24).[15] Nevertheless, the practice of converting *hukou* through marriage is strictly controlled. A local *hukou* is not guaranteed by simply marrying a local resident. Given that the *hukou* system is imposed by the Chinese government, changes in the *hukou* inheritance law likely affect status transmission exogenously in a significant manner. Because children are normally considered a valuable output of marriage, the legislation reform impacts the value of certain *hukou* types in the marriage market to some extent, as indicated by the children's future *hukou* status. The national policy reform of 1998 thus allows researchers to investigate the causal impact of the *hukou* inheritance rule on

[14] It is likely that each province implemented the law at a different time after the national law was announced. However, to the best of my knowledge, the majority of provinces implemented the law by the end of 1998, and they strictly adhered to setting 1998 as the cut-off birth year for affected newborns.

[15] As documented in Han et al. (2015) and Han and Shi (2019), the initial restriction requiring a child's *hukou* to follow that of his or her mother was established to limit migration through marriage.

marriage decisions.

2.1.3 Features of the Chinese Marriage Market

The first unique feature of the Chinese marriage market is its skewed sex ratio, i.e., the relatively high number of males relative to females. Examining this data by rural and urban areas, as shown in Figure D1, reveals a more nuanced sex ratio imbalance in rural versus urban China. Along with the segregation imposed by the *hukou* system, these two driving forces have led to an unintended, long-lasting consequence: the "bride drain" effect for rural men (Meng and Zhao, 2019).

Incorporating search frictions when interpreting marriage matching outcomes in China is crucial for two reasons. First, the Chinese marriage market is distinct from that of most Western countries because Chinese parents play a non-negligible role in their children's marriage decisions, although they are less influential nowadays.[16] In Shanghai, hundreds of parents gather at what is widely known as the *"marriage corner"* or *"local marriage market"* in People's Park every weekend to display their children's profiles on posters, in hopes of finding them an ideal match.[17] This phenomenon supports the idea of introducing *"frictions"* to the matching theory framework to allow for mismatches and non-matches in equilibrium.

Second, as a result of the social segregation imposed by the *hukou* system, rural populations, particularly rural men, face more severe friction in the marriage market. Even among rural men who migrate to urban areas to improve their income, a majority end up working in labor-intensive industries with few female workers. This makes it difficult for them to meet and interact with the urban population. Consequently, these rural migrants remain excluded from the middle class (Rozelle and Boswell, 2021). Moreover, rural men face worse marriage market conditions than rural women due to the skewed sex ratio. Despite the randomness in the searching process, marital searches tend to be targeted in that individuals use observable characteristics to decide on meetings (Ong and Wang, 2015; Ong et al., 2020). Such a targeted search is well captured in a directed search framework.

2.2 Data Source and Sample Construction

This study draws on two primary datasets: the 0.1-percent sample from the 2005 Population Census (a.k.a. the 2005 *Mini-Census*) and the 2000 *Census* published by the National Bureau of Statistics of China. Both datasets are nationally representative, covering 31 provinces, municipalities, and autonomous regions. The administrative *Mini-Census* randomly selects 0.1% of the households in China to take a survey with detailed demographic information. It collects

[16] Figure D2 shows that over 75% of the surveyed sample in the 2017 Chinese General Social Survey agreed that "parents have a say in marriage decisions."

[17] For more details, refer to https://www.abc.net.au/news/2018-04-22/marry-first-then-fall-in-love-how-marriage-evolved-in-china/9641958.

information at the individual level, including marital status, living arrangements, year of first marriage, *hukou* status (urban or rural), place of residence, place of household registration, and other basic characteristics, such as educational attainment, ethnicity, and employment status. In other words, the individual-level *Mini-Census* data provide information on specific matches, i.e., who marries whom. I also use the publicly available aggregate-level *Census* data to calculate the sex ratios for certain types of marriage markets. The variations in the sex ratio are critical for identifying men's and women's preferences in marriage decisions, since it serves as the driving factor of the equilibrium outcome in the marriage market.

For the empirical analysis, I match husbands and wives based on the reported relationships of the household head. I only retain the household heads and their spouses to avoid potential confounding in marriage decisions due to interactions between family members.[18] I focus on current marriages (97% of the reported marriages) and restrict the couples to those who got married at or above the legal marriage age in China, which is 20 for women and 22 for men. I also restrict my analysis to couples who married after 1990 to obtain a balanced sample for the policy analysis. To construct a sample of singles, which is necessary for estimating the structural model to identify the search frictions in the marriage market, I restrict the age range to individuals between 20 and 40 years old.[19] The resulting dataset comprises 135,343 observations, of which 82,679 are married couples.

The marriage market in this study is defined as the product of two characteristics: *hukou* status and educational attainment. *Hukou* status is either urban or rural, and I categorize individuals into three groups based on their educational attainment: *low education* (primary school graduates or below), *middle education* (middle school graduates), and *high education* (high school graduates or above). This division is motivated by statistical evidence from the sample distribution that the majority of people have at least a middle school diploma, as shown in Table 1.[20] Since more than 90% of marriages take place within one province, it is reasonable to consider each province as an independent marriage market. As a result, I have a total of $6 \times 6 \times 28$ marriage markets.[21] I rely on the 2000 *Census* to construct the relevant sex ratios for structural

[18] Modeling marriage decisions for couples in extended families is challenging due to more confounding factors resulting from the potential correlations or interactions within the family. Additionally, the decision to get married may vary between household heads and non-household heads, with the latter being influenced by factors like housing. I leave the incorporation of living arrangement decisions or housing market factors in the current marriage search and matching model for future studies. Nonetheless, the DID analysis that employs all couples within the specified time frame yields robust results, as demonstrated in Appendix B.

[19] This age range accommodates the majority of samples from the *Mini-Census*, given that by the age of 35, approximately 90% of the population, irrespective of gender, is married, as illustrated by Figures D3 and D4 in the Appendix. I additionally exclude observations from the three provinces with the smallest populations (less than 5% of the Census population) to ensure convergence: *Tibet, Qinghai, and Ningxia*.

[20] According to the 2005 *Mini-Census*, the proportion of individuals who received a college degree was just 6.22% in China overall. A closer inspection of the data with respect to *hukou* status reveals an exceptionally low prevalence of college graduates among rural populations, around 0.62%. The worse educational outcomes in rural China have been documented in previous studies such as Zhang et al. (2002) and Zhang and Kanbur (2005). Therefore, for the analysis, a high school education can be considered an appropriate proxy for the highly-educated group.

[21] The first six corresponds to $6 = 3 \times 2$ types of men (three education groups and two *hukou* types). Similarly, the second

estimation.[22]

Table 1: Summary Statistics for Married Couples

	N	Mean	Median	se(Mean)
𝟙(Wife urban hukou)	82,679	0.357	0	0.002
𝟙(Husband urban hukou)	82,679	0.341	0	0.002
Difference in years of schooling (husband-wife)	82,679	0.639	0	0.008
Age gap (husband-wife)	82,679	1.464	1	0.008
𝟙(Wife lower then primary school)	82,679	0.028	0	0.001
𝟙(Wife primary school)	82,679	0.214	0	0.001
𝟙(Wife middle school)	82,679	0.489	0	0.002
𝟙(Wife high school)	82,679	0.148	0	0.001
𝟙(Wife college or above)	82,679	0.120	0	0.001
𝟙(Husband lower than primary school)	82,679	0.008	0	0.000
𝟙(Husband primary school)	82,679	0.150	0	0.001
𝟙(Husband middle school)	82,679	0.518	1	0.002
𝟙(Husband high school)	82,679	0.176	0	0.001
𝟙(Husband college or above)	82,679	0.148	0	0.001
Wife marriage age	82,679	23.268	23	0.008
Husband marriage age	82,679	24.732	24	0.009

Note: This table reports summary statistics for the wife and the husband's characteristics for married couples. The sample is restricted to couples married above legal marriage ages (22 for male and 20 for female) and those married after 1990. The data is from 2005 *Mini-Census*.

2.3 Motivating Evidence

Figure 1 depicts the matching pattern among married couples in the analysis sample. Each cell in the graph represents the proportion of matches for that specific type of individual. The cells along the diagonal have the darkest colors, suggesting that the vast majority (over 85%) are same-type matches. Exploring the matches in detail reveals a number of important findings. First, intermarriage between rural women and urban men is more prevalent than intermarriage between urban women and rural men. This pattern is also consistent with Figure 2, which shows the trend of intermarriages by *hukou* status. Second, as shown in Table D1, the assortativeness of matching on the dimension of *hukou* status is greater than on the dimension of educational attainment. This finding implies that the *hukou* policy has exacerbated inequality in the Chinese marriage market by generating social segregation between the two groups of individuals. Table D1 also shows that women are more likely to "marry down" on education while "marrying up" on *hukou*.[23] This suggests that multi-dimensional attributes are valued in

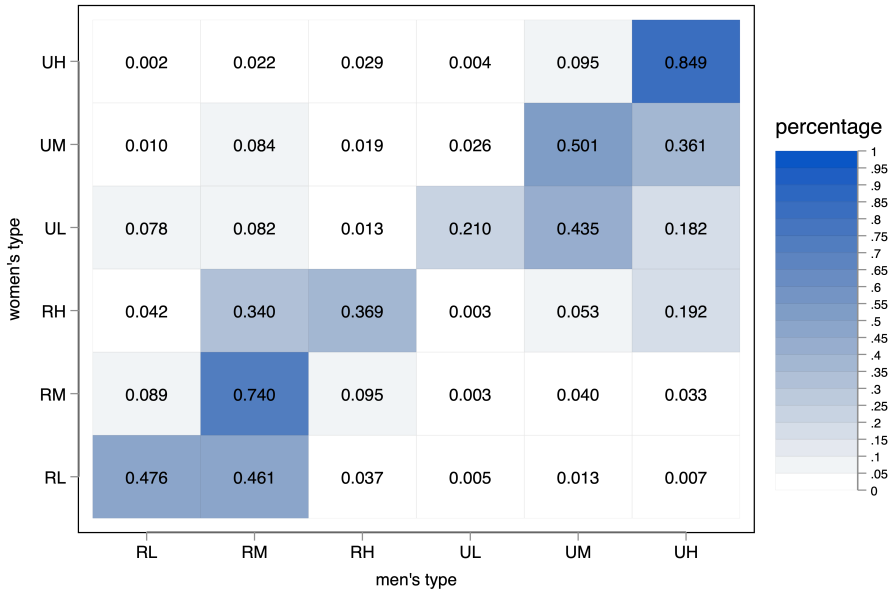
six corresponds to 6 types of women. The number 28 denotes the number of provinces in the estimation sample.

[22]The detailed construction method can be found in Appendix A.

[23]The term "marry-up" ("marry-down") refers to a situation in which the matched spouse possesses a more (less) desirable trait than the individual.

the marriage market, and that there is likely a trade-off between social status and educational attainment in marriage decisions. The dissimilar pattern on the two dimensions also implies that it is necessary to evaluate both attributes jointly when examining individuals' preferences for partners in marriage decisions. Table 1 delves deeper into the summary statistics of married couples' characteristics. We can see that husbands tend to have a higher level of education than wives. The age gap between husband and wife is roughly two years on average, and the average age of marriage for women is around 23, while for men, it is around 25.

Figure 1: The Match Distribution



Note: This figure visualizes the match distribution of the estimation sample. The first letter denotes rural or urban *hukou* : R for rural and U for urban. The second letter denotes the educational category: L for low education, M for middle education, and H for high education. The number in each cell denotes the proportion of couples of this specific type of match in the estimation sample. For each row, the numbers add up to 1. The x-axis denotes men's type while the y-axis denotes women's type. The data is from 2005 *Mini-Census*.

2.3.1 The Effect of *Hukou* Policy on Marriage Outcomes

In this section, I provide suggestive evidence on how a status-related policy affects marriage outcomes via the difference-in-differences (DID) approach. The analysis builds on the 1998 national reform, which grants men the same rights as women to pass on *hukou* to their children. Before August 1998, children in China could only inherit *hukou* status from their mothers (Han et al., 2015; Han and Shi, 2019; Nie and Xing, 2011).

I use the DID analysis to capture the effect of the inheritance rule change. This analysis serves as a suggestive presentation of trends in marriage outcomes following the policy, which I will further explore using the structural model. The first layer of difference lies in whether a

given man was younger than 22, the legal marriage age, in 1998. I construct the second source of variation based on the absolute value of the difference between urban and rural population shares for each province in 1998, $|PopShare_{urban} - PopShare_{rural}|$. I obtain the data from the *1998 China Population and Employment Statistics Yearbook*. The second difference is valid since a larger value of the measure implies a more imbalanced urban and rural population. As a result of the policy, we would expect to see a smaller proportion of intermarriages between urban and rural *hukou* holders, given that each province can be viewed as a separate marriage market. Figure D5 visualizes the variations in the urban and rural population gap for each province. The identification assumption of the DID approach is that without the 1998 policy reform, a common trend would exist in intermarriage rates across provinces with different levels of rural and urban population share gap. Figure D6 visualizes the parallel trend test results. Specifically, the empirical model follows:

$$Y_{ic} = \beta_0 + \beta_1 \mathbb{1}(Age_{ic}^{1998} < 22) + \beta_2 \mathbb{1}(Age_{ic}^{1998} < 22) \times G_c + \mathbf{X}_{ic}\boldsymbol{\mu} + \lambda_c + \varepsilon_{ic}, \tag{1}$$

where Y_{ic} is the indicator of intermarriage across *hukou* types for a man i in province c, which takes the value 1 when a husband and wife have different *hukou* types,[24] Age_{ic} is the man i's age in 1998; G_c is the absolute value of the difference between the urban and rural population shares in province c in 1998, which captures the scope of the policy impact on intermarriages; and $\mathbf{X}_{ic}$ is a vector of covariates including individual characteristics, such as educational attainment, marriage age, and whether the man belongs to a minority group. I also include province fixed effects, λ_c, which remove all time-invariant province-specific heterogeneity that is correlated with the other covariates and the outcome variable. I estimate (1) using the 2005 *Mini-Census* data. The sample is restricted to couples married after 1990. I would expect the parameter of interest β_2 to be negative if people indeed responded to the policy, as a larger G_c implies a more imbalanced urban and rural population, leading to a smaller likelihood of intermarriage in that province. In Appendix B, I also try alternative specifications to show the robustness of the results.

Figure 2 depicts the motivating fact that intermarriages between rural and urban *hukou* gradually increased. More importantly, starting in 1998, the intermarriages between rural women and urban men experienced a larger increase compared to the intermarriages between urban women and rural men. Table 2 shows the results of the DID analysis. The odd columns report results with no individual controls added, whereas the even columns report results with additional individual controls. We can see that β_2 is negative in all specifications. Given how

[24] Ideally, we should use the *hukou* information at the time of marriage for a more accurate measure. However, only data from the 2006 Chinese General Social Survey (CGSS) has this information. The data show that among married respondents, about 16.88% acquired urban *hukou* after marriage. However, the reasons for a *hukou* status change can be multi-fold. Among them, job promotion is the major reason. Only around 7% of the population changed their *hukou* status due to marriage. For more information about CGSS data, please refer to http://cgss.ruc.edu.cn/English/Home.htm.

Figure 2: Trend of Intermarriage

Note: This figure plots the trend of intermarriage between urban and rural *hukou*. The solid line refers to the intermarriage between urban men and rural women, while the dashed line refers to the intermarriage between rural men and urban women. The vertical line refers to the year of the policy change - 1998. The data is from 2005 *Mini-Census*.

I constructed G_c, this suggests an increase of intermarriages between different *hukou* holders. When looking at specific types of intermarriages, I find a more substantial increase in magnitude for intermarriages between rural women and urban men, compared to intermarriages between rural men and urban women.[25] This difference in magnitude remains noteworthy and statistically significant even when considering any controlling factors. Though the 1998 *hukou* policy altered the inheritance law rather than directly removing status differences, the DID analysis shows that status-related policies can significantly affect men's and women's marriage decisions. Since individuals care about their offspring's *hukou* status, that would affect how they select their potential spouses, thus changing the value of *hukou* status and reshaping the marriage market.

3 A Model of Marriage Search and Matching

Research has shown that when inequality increases, intermarriage decreases as it becomes more costly to marry someone from a lower social class (Fernandez et al., 2005). However, the results of the data exploration exercise presented in the previous section suggest the opposite may be true for intermarriage among different *hukou* categories in China, where there are now in-

[25] There is an increase in *hukou* intermarriage among rural men after the policy reform, possibly due to the migration flows from rural to urban areas.

Table 2: The Impact of *Hukou* Policy on Intermarriage

	(1) intermarriage	(2) intermarriage	(3) rural w -urban m	(4) rural w-urban m	(5) rural m-urban w	(6) rural m-urban w
$\mathbb{1}(Age^{1998} < 22)$	0.010	0.030***	0.007	0.025***	0.003	0.006
	(0.009)	(0.009)	(0.008)	(0.007)	(0.004)	(0.005)
$\mathbb{1}(Age^{1998} < 22) \times G_c$	-0.049**	-0.054***	-0.036**	-0.040**	-0.013	-0.014
	(0.015)	(0.015)	(0.013)	(0.012)	(0.008)	(0.008)
Individual controls	No	Yes	No	Yes	No	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
p-value for (3)=(5): 0.094						
p-value for (4)=(6): 0.070						
Mean Dep. Var.	0.065	0.065	0.040	0.040	0.024	0.024
R^2	0.016	0.020	0.011	0.018	0.006	0.006
N	82,679	82,679	82,679	82,679	82,679	82,679

Note: Intermarriage is defined as marriages between urban and rural *hukou*. G_c is the absolute value of the difference between urban and rural population share in province c in 1998. Individual controls include the man's marriage age, education level, and whether belongs to a minority group. I restrict samples to those who got married after 1990 for all the specifications. I also test the equality of β_2 for both types of intermarriages and report the corresponding p-values. Robust standard errors, clustered at the city level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$. *Data Source:* 2005 *Mini-Census*.

creased incentives for such marriages following the policy reform. Additionally, the increase in intermarriage between urban and rural *hukou* holders is not the same for both genders, as rural women are more likely to marry urban men than rural men are to marry urban women. Thus, it is crucial to separately identify men's and women's preferences for *hukou* status in marriage decisions to gain a complete understanding of the policy impact. However, the reduced-form analysis is limited in obtaining separate individual preferences for men and women. It is also silent in capturing the general equilibrium effect of the policy due to simultaneous changes in the sex ratio in the marriage market. I thus proceed to develop and estimate a structural model of marriage search and matching. With a structural model, I am also able to conduct counterfactual policy analyses.

3.1 Model Set-up

Following Arcidiacono et al. (2016), I construct a two-sided directed search and matching model with nontransferable utility and consider only opposite-sex, one-to-one matching.[26] I categorize each male as a type m, where $m \in \{1, 2, \ldots, M\}$. Similarly, each woman is given a type w, where $w \in \{1, 2, \ldots, W\}$.[27] For males (females) there are then $W(M)$ types of mates. Let im indicate the ith member of type m. In the market, agents search for specific characteristics. The search decision follows a discrete choice framework.

As it is a directed search process, the expected utility from searching in a particular market

[26]The assumption of transferable utility is more realistic in the case of marriage. However, as noted in Chiappori and Salanié (2021), in societies ruled by very rigid social norms, transfers may be constrained and frictions can be critical. When emphasizing the role of frictions, uncertainty about matching serves to coordinate agents' decisions and, therefore, to clear the market instead of transferring utility. In most transferable utility models, transfers are symmetric, but in my search model, the transfer is essentially the disutility associated with the possibility of not matching. Furthermore, in a framework of transferrable utility, it is unable to identify women's and men's preferences over spousal traits separately, as its main goal is to identify the joint marital surplus or gains from marriage.

[27]An individual's type denotes a collection of observed characteristics such as age, educational attainment, or attractiveness.

is then the probability of matching in the market times the utility conditional on matching. Specifically, the expected utility for a m-type man searching for a w-type woman is modeled as:

$$E\left(U_{im}^{w}\right) = P_{m}^{w} \cdot e^{\mu_{m}^{w}+\varepsilon_{im}^{w}}, \tag{2}$$

where P_m^w is the probability of matching in the marriage market composed of type w women for men of type m, μ_m^w is the deterministic utility for this marriage market for a m-type man, and ε_{im}^w is an individual specific error term. It is known to the individual before the search decision, but unknown to the analyst.

After taking the logarithm and following the utility maximization rule, individual i of type m then chooses to search for a woman of type w, $d_{im} = \{w\}$ when:

$$\{w\} = \arg\max_{w'} \mu_{m}^{w'} + \ln\left(P_{m}^{w'}\right) + \varepsilon_{im}^{w'}. \tag{3}$$

Under the standard assumption that the ε_{im}^w's are independent and identically distributed (i.i.d.) type I extreme value errors with the error variance σ, the probability of a m-type man searching for a w-type woman, ψ_m^w follows:

$$\Pr(w|m) = \psi_{m}^{w} = \frac{\exp\left(\frac{\mu_{m}^{w}+\ln(P_{m}^{w})}{\sigma}\right)}{\sum_{w'}\exp\left(\frac{\mu_{m}^{w'}+\ln\left(P_{m}^{w'}\right)}{\sigma}\right)}. \tag{4}$$

As for the matching process, it is modeled as a Cobb-Douglas production function, taking as inputs the number of searching men and the number of searching women in each market and giving the number of matches in each market as an output.[28]

The number of matches in market $\{m,w\}$ is modeled as:

$$X_{mw} = A(\psi_{m}^{w}N_{m})^{\rho}\left(\psi_{w}^{m}N_{w}\right)^{1-\rho}, \tag{5}$$

where A captures the search friction, ρ determines the elasticity of substitution, $\psi_m^w N_m$ is the number of men of type m searching for women of type w, similarly for $\psi_w^m N_w$. Note that N_m and N_w represent the number of m-type men and of w-type women overall.

Under the model setup, utility is type specific instead of individual specific. Note that the probability of matching and the deterministic utility from matching vary only at the type level rather than at the individual level. Stating it differently, all men of type m searching for a spouse of type w have the same probabilities of matching and the same deterministic components of

[28]The model is assumed to follow the Cobb-Douglas setting, which can be nested into the more general CES production function setting with the elasticity of substitution converging to 0.

utility. This also means that the probability of matching is only affected by the individual's and partner's types, and not by any idiosyncratic individual trait. The only individual-specific element of the expected utility is the logit error. Thus, all m-type men searching in the same market have the same probabilities of matching, P_m^w is given by[29]

$$\begin{aligned} P_m^w &= \frac{X_{mw}}{\psi_m^w N_m} = \frac{A(\psi_m^w N_m)^\rho (\psi_w^m N_w)^{1-\rho}}{\psi_m^w N_m} \\ &= A\left(\frac{\psi_w^m N_w}{\psi_m^w N_m}\right)^{1-\rho}. \end{aligned} \tag{6}$$

Here, the ratio of N_w by N_m represents the gender ratio (i.e., the number of women relative to men). This equation captures how the gender ratio affects search decisions, which further shifts the equilibrium outcome depending on ρ.

The equilibrium is defined as the fixed point solution of the searching probability. From (4) and (6), the searching probability ψ and the matching probability P are correlated with each other. With the specific distribution of the unobservable preference parameter, the fixed point equation follows:

$$\psi_m^w = \frac{\exp\left(\mu_m^w + \ln\left[P_m^w\left(\psi_w^m, \psi_m^w\right)\right]\right)}{\sum_{w'} \exp\left(\mu_m^{w'} + \ln\left[P_m^{w'}\left(\psi_w^{m'}, \psi_{m'}^w\right)\right]\right)}. \tag{7}$$

Moreover, equilibrium should also include the constraint that the market shares must sum to 1 for both men and women. Equilibrium in this model is then characterized by stacking the $(M-1)$ and $(W-1)$ shares and solving for the fixed point by the set of equations (7) for men and the analogous set for women. The searching probability ψ is a continuous mapping on a compact and convex space. The Brouwer's Fixed Point Theorem guarantees the existence and uniqueness of the equilibrium.

3.2 Identification and Estimation

In this subsection, I first present the model specifications, with a focus on the matching technology A and the deterministic utility of matching types μ. I then discuss the identification of the parameters and the estimation strategy.

Motivated by the DID analysis results, I specify the matching technology A to capture the 1998 policy change, given by:

$$A = A_0 + A_1 \mathbb{1}(Age_{1998m} < 22) + A_2 \mathbb{1}(PH_m = 1) + A_3 \mathbb{1}(Age_{1998m} < 22)\mathbb{1}(PH_m = 1),$$

where $\mathbb{1}(Age_{1998m} < 22)$ is the indicator of whether the man was younger than 22 in 1998, $\mathbb{1}(PH_m = 1)$ is the indicator of the man being rural *hukou*, A_0, A_1, A_2, A_3 are the parameters in the matching technology function. The specification aims to capture the heterogeneous effects

[29] The probability of matching for w-type women is derived to follow a symmetric form.

of the policy change on rural men compared to urban men, motivated by the results in Section 2.3.1.[30]

Estimation results reported in the next section follow the specification of deterministic utility as described below. In the baseline model, the deterministic utility for men and women is given by:

$$\mu_{mw} = \alpha_1 SE_{mw} + \alpha_2 \mathbb{1}(PE_w = 2) + \alpha_3 \mathbb{1}(PE_w = 3) + \alpha_4 \mathbb{1}(PE_w > PE_m) + \alpha_5 SH_{mw} + \alpha_6 \mathbb{1}(PH_w = 1)$$
$$\mu_{wm} = \alpha_7 SE_{wm} + \alpha_8 \mathbb{1}(PE_m = 2) + \alpha_9 \mathbb{1}(PE_m = 3) + \alpha_{10} \mathbb{1}(PE_m > PE_w) + \alpha_{11} SH_{wm} + \alpha_{12} \mathbb{1}(PH_m = 1),$$

where the intercept is normalized to 0. Denote the educational group associated with a m-type man as $E_m \in \{1,2,3\}$. When a man searches for a w-type woman, the educational group of the partner is PE_w. The term SE_{mw} is an indicator function taking value 1 when wife and husband belong to the same educational group, similarly for SE_{wm}. I also specify the deterministic utility to capture the differential preferences over relative educational attainment by gender. Denoting $H_m \in \{1,2\}$ as men's *hukou* status, with 1 denoting rural *hukou* and 2 denoting urban *hukou*, then $SH_{mw} = I(H_m = PH_w)$ represents whether husband and wife have the same *hukou* status. I allow for the effect of spousal educational attainment and *hukou* status to vary by gender. For a richer model, I add the interaction of educational group and *hukou* type to account for potential heterogeneity in preferences.

In total, there are two sets of parameters to be estimated: those of the utility function, α, and those of the matching function, ρ and A_0, A_1, A_2, A_3. Denote θ as the set $\{\alpha, \rho, A_0, A_1, A_2, A_3, \sigma\}$. The likelihood for the ith man of type m can be formulated as:

$$\begin{aligned} L_{im}(\theta) = &\left[\prod_w \left([\psi_m^w(\theta)] \times [P_m^w(\theta)]\right)^{I(d_{im}=\{w\})}\right]^{I(y_{im}=1)} \\ &\times \left[\sum_w \psi_m^w(\theta) \times (1 - P_m^w(\theta))\right]^{I(y_{im}=0)}, \end{aligned} \tag{8}$$

where $y_{im} = 1$ denotes the ith man is matched and 0 otherwise. The man's search decision d_{im} is only observed when matched. Thus, the search decision for those who are not matched needs to be integrated out. The numerical integration is the same as summing over all the possible markets defined by type w. Here, as the equilibrium is captured by the fixed point equation, the equilibrium searching probability ψ_m^w is a function of the matching probabilities (P_m^w, P_w^m).

[30] The rural and urban population gap can only capture the effects of the policy change on the rate of intermarriages, but limited in explaining the heterogeneous effects on the two types of intermarriages.

The log likelihood for the ith man of type m follows:

$$\begin{aligned} LL_{im}(\theta) =& I(y_{im}=1)\left[\sum_w I(d_{im}=\{w\})\left(\ln\left[\psi_m^w(\theta)\right]+\ln\left[P_m^w(\theta)\right]\right)\right] \\ &+ I(y_{im}=0)\ln\left[\sum_w \psi_m^w(\theta)\times\left[1-P_m^w(\theta)\right]\right]. \end{aligned} \tag{9}$$

The marriage market is defined within each province in China and I assume that the provinces are independent of each other. Thus, I need to sum the log likelihoods over all the possible male types at each province, denoted as $c \in \{1,\cdots,C\}$. The estimated parameters are obtained as:

$$\hat{\theta} = \arg\max_{\theta}\left(\sum_c\sum_m\sum_{i=1}^{N_m^c} LL_{im}^c(\theta)\right). \tag{10}$$

At each iteration, the fixed point equation for the searching probabilities expressed by (7) is solved. In sum, within each province and for each iteration of the likelihood function, I must first solve an $m \times w$ fixed point in the search probabilities $[\psi_w^m, \psi_m^w]$ at the type level.

I hereby provide a brief heuristic discussion on how researchers can identify the parameters $\theta = \{\alpha, \rho, A_0, A_1, A_2, A_3, \sigma\}$.[31] The identification of the preference parameters α follows from identification in a traditional discrete choice model, by covariation between the observed choices and individual and partner characteristics, which relies on the variations of the sex ratios on each marriage market defined by the cross-product of characteristics. To separately identify preferences from both sides of the market (men and women) requires variations in the choice set. The variations come from estimating the model across provinces, exploiting the marriage markets' geographic segmentation, which generates different numbers of singles across markets. The parameters of the matching function are identified by the variations in matches across the 28 provinces with different sex ratios and overall match rates. The key parameter A_3 that captures the heterogeneous effects of the policy on the marriage market can be identified by the exogenous variation of the 1998 policy.[32]

I solve the model by maximum likelihood estimation (MLE). Particularly, the constrained optimization is formulated as:

$$\begin{aligned} &\max_\theta \textstyle\sum_c\sum_m\sum_{i=1}^{N_m^c} LL_{im}^c(\theta) \\ \text{subject to}\quad &\psi_m^w = \frac{\exp\left(\frac{\mu_m^w + \ln\left[P_m^w(\psi_w^m,\psi_m^w)\right]}{\sigma}\right)}{\sum_{w'}\exp\left(\frac{\mu_m^{w'} + \ln\left[P_m^{w'}\left(\psi_{w'}^m,\psi_m^{w'}\right)\right]}{\sigma}\right)}. \end{aligned} \tag{11}$$

The objective function is the log-likelihood following (9). The fixed point formulas of

[31] I also summarize the identification strategy for each parameter in Table D2. For a more detailed discussion on the identification of similar models, one can refer to Hsieh (2012).

[32] In Appendix C, I also provide examples of identification on several parameters with graph illustrations.

searching and matching probabilities on each market are imposed as the nonlinear constraints to the optimization problem. The searching probabilities are represented by ψ, while the matching probabilities are represented by P. The corresponding men's and women's types are denoted by $\{m, w\}$. The deterministic utility component is denoted by μ, while the variance of the error term is denoted by σ^2. In estimation, I normalize σ to be equal to one, following the traditional practice of the standard discrete choice logit model. In contrast to Arcidiacono et al. (2016) who use the Nested Fixed-Point (NFXP) algorithm to solve the optimization problem, I follow Dubé et al. (2012) to implement the Mathematical Programming with Equilibrium Constraints (MPEC) method. The MPEC method has been proven to be more robust and efficient than NFXP for estimating empirical matching models (Dong et al., 2020).

4 Estimation Results and Counterfactual Analysis

4.1 Estimation Results

Tables 3 and 4 display the estimation results of the baseline model (i) and the richer model (ii) that accounts for the interaction of educational attainment and status. The standard errors are obtained via 200 bootstrap simulations. First, for the preference parameters, I find that both men and women prefer high educational attainment. Interestingly, women prefer their husbands to have higher educational attainment, while men dislike women with higher educational attainment. The finding is consistent with patterns from Beauchamp et al. (2021), who estimated matching preferences in the Indian marriage markets. The divergent education preferences between men and women may be linked to the heterogeneity observed in the proportions of single individuals by gender and level of educational attainment, as presented in Figure D7.[33] This suggests that men and women value education differently in the marriage market, which has implications for individual educational investment decisions and potential gender disparities in educational outcomes. The relationship between education and marital outcomes, particularly with regard to gender, remains an important area of inquiry for both theoretical and practical reasons.

In terms of the heterogeneity across *hukou* types, compared to rural women, urban women have a stronger preference for having the same educational attainment as their spouses. In contrast, men exhibit a different pattern— the preference for having the same educational attainment as one's spouse is stronger for rural men than for urban men. When it comes to *hukou* status, I find that women have a preference for having the same *hukou* as their spouses, whereas men do not. Urban men prefer a rural status whereas urban women dislike it, indicating that women tend to marry up in *hukou* status while men tend to marry down, especially when there

[33] Figure D7 displays the singleness rates among individuals aged 35 and above, separately by gender and education level. As shown in the figure, women with higher education have a slightly higher likelihood of remaining single, while men do not exhibit the same pattern.

is a shortage of women in the market. I also find that both rural women and men prefer to have the same status as their partners.

For the matching technology parameters, both specifications yield a significant and negative estimate for A_3, indicating that rural men, in comparison to urban men, indeed experienced higher friction in marriage markets after the policy. This outcome can be attributed to the fact that the pool of potential spouses for rural men diminished after the reform, as more rural women are likely to be matched with urban men, as can be observed from the suggestive evidence in Section 2.3.1. Consequently, rural men encounter greater impediments due to their reduced marriage prospects. Figure D8 compares the observed married rates and single rates for the six types of individuals with the predicted probabilities in equilibrium using the richer model. This figure shows that the model provides a reasonable fit to the underlying data. As a result, I mainly use the richer model as the preferred specification.

4.2 Counterfactual Analysis

Given the model estimates, I now examine a series of counterfactual environments to see how women's and men's choices and matches respond in equilibrium. First, I explore how the marriage market's constitution, particularly the *hukou* intermarriage rates and marriage matching pattern, would be reshaped had the 1998 policy not been implemented, by setting the policy indicator to 0 in the matching technology specification. Second, I investigate the welfare consequences of a set of counterfactual policies. The first counterfactual scenario simulates an education policy aiming to equate the distribution of educational attainment between rural and urban populations for women and men, respectively. The second one predicts the new marriage matching pattern after varying the sex ratio for each marriage market. Due to the recent relaxation of the fertility restriction, the sex ratio in the marriage market has become gradually stable, which could further affect matching outcomes. The third counterfactual experiment predicts the matching pattern after completely removing the *hukou* differences. This exercise has implications for understanding the effect of the 2014 national *hukou* reform on the marriage market.

4.2.1 Counterfactual Simulation in the Absence of the 1998 Policy

The first counterfactual policy experiment simulates the marriage market had the 1998 policy not been implemented. Figure 3 shows the change in marriage matching patterns for the six types of individuals defined by the cross product of educational attainment and *hukou* status. Figure 3(a) depicts the baseline matching pattern along the two dimensions from the data. We can see that most marriages happen along the same attributes for men and women, while a majority of the intermarriages between rural men and urban women occur for high-education

Table 3: Female vs. Male Preferences Estimates

	(i) baseline	(ii) with interaction
Female Preferences α_w		
spouse has higher education	0.360	0.228
	(0.010)	(0.007)
same education group	0.338	0.303
	(0.003)	(0.004)
spouse is middle school graduate	-0.012	0.057
	(0.005)	(0.006)
spouse is at least high school graduate	-0.206	0.003
	(0.005)	(0.007)
$\mathbb{1}$(rural hukou) x same education group		-0.052
		(0.005)
$\mathbb{1}$(rural hukou) x middle school		-0.069
		(0.006)
$\mathbb{1}$(rural hukou) x at least high school		-0.213
		(0.006)
same hukou	0.269	0.098
	(0.075)	(0.004)
$\mathbb{1}$(rural hukou) x same hukou		0.407
		(0.007)
spouse is rural hukou	-0.123	-0.270
	(0.075)	(0.004)
Male Preferences α_m		
spouse has higher education	0.088	-0.069
	(0.015)	(0.017)
same education group	0.182	0.051
	(0.014)	(0.014)
spouse is middle school graduate	-0.098	0.084
	(0.006)	(0.014)
spouse is at least high school graduate	-0.167	0.008
	(0.007)	(0.017)
$\mathbb{1}$(rural hukou) x same education group		0.035
		(0.011)
$\mathbb{1}$(rural hukou) x middle school		-0.190
		(0.015)
$\mathbb{1}$(rural hukou) x at least high school		-0.197
		(0.017)
same hukou	0.173	-0.059
	(0.081)	(0.011)
$\mathbb{1}$(rural hukou) x same hukou		0.139
		(0.011)
spouse is rural hukou	0.128	0.067
	(0.081)	(0.012)

Note: This table reports model estimates for preference parameters for males and females, respectively. The standard errors are obtained via 200 bootstraps. *Data Source:* 2000 *Census* and 2005 *Mini-Census*.

Table 4: Matching Parameter Estimates

Matching Parameters	(i) baseline	(ii) with interaction
A_0	0.785	0.820
	(0.001)	(0.002)
A_1	-0.576	-0.607
	(0.002)	(0.002)
A_2	0.067	0.043
	(0.001)	(0.002)
A_3	-0.054	-0.038
	(0.002)	(0.003)
ρ	0.308	0.217
	(0.008)	(0.006)
N	135,343	135,343
log(likelihood)	-132077.727	-130745.479

Note: This table reports model estimates for the matching technology specification. I normalize σ to be equal to one, following the traditional practice of the standard discrete choice logit model. The standard errors are obtained via 200 bootstraps. *Data Source:* 2000 *Census* and 2005 *Mini-Census*.

men. As shown in Figure 3(b), in the absence of the 1998 policy, the intermarriages between urban men and rural women would decrease, while marriages between rural men and rural women would increase.

I further quantify the welfare change for different types of individuals based on utility changes before and after the counterfactual policy. Each change in utility is calculated based on the logit-consumer surplus functional form.[34] I define an individual as a "winner" when the corresponding utility change is positive and vice versa.

Figure 4 displays the results for welfare changes. I find that in the absence of the 1998 policy, all types of women benefit, urban and low-educated men experience the greatest loss, and rural and high-educated men gain the most. This, in turn, suggests that both rural and urban women would be worse off due to this policy change. Their worse outcomes may be attributed to several reasons. First, according to the model estimates in Table 3, rural women prefer to have the same *hukou* as their spouses. Since the policy increased intermarriages, especially intermarriages between urban men and rural women, it leads to potentially worse outcomes for rural women. Rural women may also be harmed since the policy also increased friction between rural women and rural men. My finding of rural women being worse off is also

[34] According to Train (2009), the welfare change can be calculated as the change in consumer surplus:

$$\Delta E(CS_n) = \frac{1}{\alpha_n}[\ln(\sum_{j=1}^{J^1} \exp^{V^1_{nj}}) - \ln(\sum_{j=1}^{J^0} \exp^{V^0_{nj}})],$$

where I set $\alpha_n = 1$ in my analysis to capture the change of signs before and after the policy, J^1 is the maximum number of choice sets in the counterfactual scenario, V^1_{nj} is the value of deterministic utility for the counterfactual case, and J^0 and V^0_{nj} are the counterparts at the baseline.

consistent with previous studies showing that these "unequal" matches may bring unintended consequences, such as intimate partner violence. For instance, through qualitative interviews, Lui (2018) found that rural wives of urban men are socially marginalized and treated as "unpaid reproductive workers" by their husbands' families, leading to their exclusion from urban society.

As for urban women, the preference estimates in Table 3 indicate that they prefer urban spouses, but the policy has made it more difficult to find partners due to increased competition from rural women. Additionally, the lower level of education among the rural population is not attractive to urban women. The policy may have also had negative consequences for urban women matched with urban men, as it may have weakened their bargaining power in a tighter remarriage market (Han and Shi, 2019).

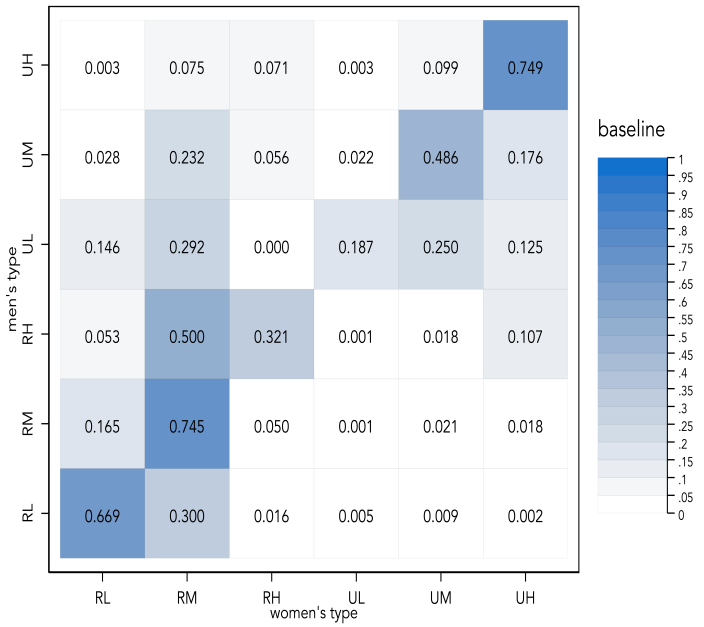
Concerning the outcomes for men, the reform has made rural men worse off, primarily due to the relatively larger friction they face in the marriage market compared to urban men. Rural residents, particularly rural men, are subject to *dual pressures*. For one thing, they greatly suffer from the extremely unbalanced sex ratio, which is related to the well-known *One Child Policy* and the deeply-rooted son preference in China (Ebenstein, 2011).[35] For another, due to the segregation imposed by the unique *hukou* system, the rural population in China faces considerable challenges in finding spouses, especially rural men. Most of the rural population works in labor-intensive industries, and it is usually extremely difficult for them to form relationships with their urban counterparts (Meng and Zhao, 2019). As a result of the policy, rural men's unsettling position may be worsened by the increased possibility of prospective rural spouses being "snatched up" by urban competitors, drastically reducing their low-friction match options. Despite an increase in intermarriages between rural men and urban women following the reform, rural men still experienced a net loss in welfare. According to the model estimates, rural men tend to prefer spouses with the same *hukou* and have a more negative view towards highly educated women. Therefore, marrying urban women makes them worse off because they have a lower chance of getting married to rural women after the policy change.

It is expected that urban low-educated men would lose the most in the absence of the reform, as most intermarriages happened between these men and rural women at the baseline. In the absence of the 1998 policy, these intermarriages would decrease, and the men's desired spouses would be matched with those rural and high-educated men, resulting in the most substantial welfare losses of all the groups.

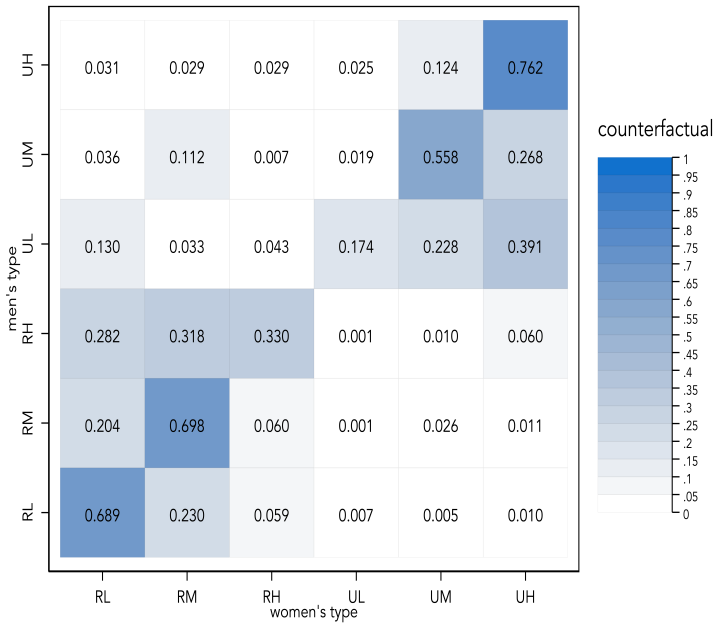
To better interpret these results, I next quantify the welfare changes by measuring changes by singleness rates, which would be directly related to the change of market friction. Specifically, I transfer the change of the policy impact on friction, measured by A_3, to the change of the proportion of single men, and make similar adjustments for single women. This is done

[35] Figure D1 demonstrates a severe sex ratio imbalance in rural versus urban China.

Figure 3: Marriage Matching Distributions With and Without the 1998 Policy



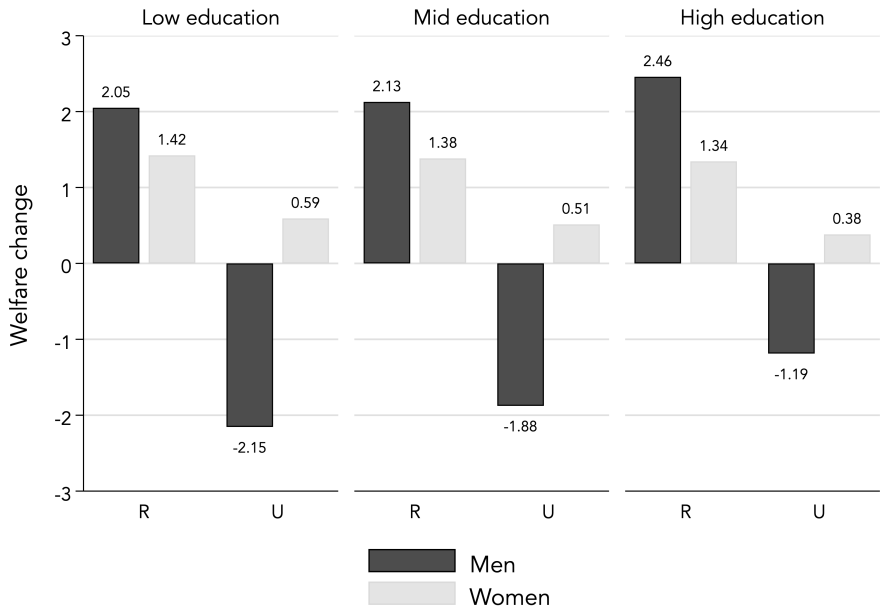
(a) Baseline matching pattern (with the 1998 policy)



(b) Without the 1998 policy

Note: The two figures show the matching distributions of the six types of individuals. Panel A shows the pattern at the baseline from the data (with the 1998 policy). Panel B shows the pattern for the counterfactual scenario (without the 1998 policy). The first letter denotes rural or urban *hukou*: R for rural and U for urban. The second letter denotes the educational category: L for low education, M for middle education, and H for high education. The number in each cell denotes the proportion of couples of this specific type of match in the estimation sample. For each row, the numbers add up to 1. The x-axis denotes men's type while the y-axis denotes women's type. Data are from 2000 *Census* and 2005 *Mini-Census*.

Figure 4: Welfare Change in the Absence of 1998 Policy



Note: This figure shows the welfare change for different types of individuals by educational attainment and *hukou* status in the absence of the 1998 policy. R/U refers to rural/urban *hukou*. Data are from 2000 *Census* and 2005 *Mini-Census*.

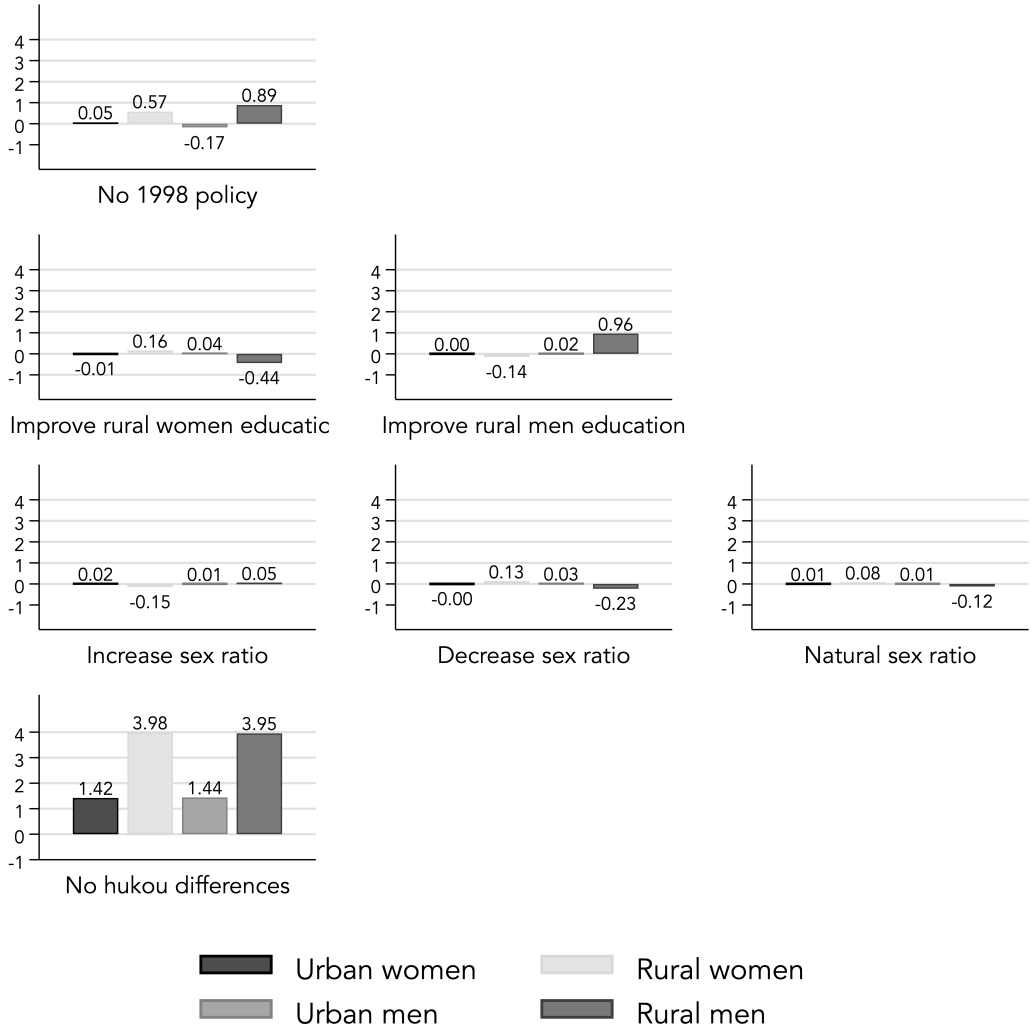
by recalculating the value of the friction parameter A_0 to let the utility at baseline ($A_3 \neq 0$) equal the utility under the counterfactual ($A_3 = 0$), analogous to measuring the "compensating variation" of the policy change. After obtaining the new value of A_0, I simulate the matching pattern in the estimation sample to extract the changes in the respective proportions of singles for men and women. Results are presented in Figure D9. We can see that in order to let low education and rural men gain a welfare of 2.05 in the absence of the 1998 policy, equivalent to a 7.2% decrease of single men in the marriage market. This illustrates the significant marriage distortions induced by the status-related policy. I also visualize the change in intermarriages between rural women and urban men for each province in Figure D10. The pattern shows that intermarriages would decrease and provinces with higher urban populations (relative to rural populations) would experience a larger decrease. In Table D3, I further show the correlation between winner proportions and province characteristics. The negative coefficient suggests that provinces with higher urban-to-rural population ratios would benefit the least from the removal of the 1998 policy.

4.2.2 Counterfactual Policy Experiments

I now turn to analyzing the effect of a series of counterfactual policies. The choice of these policies is motivated by recent developments in marriage practices and legislation in China,

which I briefly discuss below. Figure 5 presents the changes in ex-ante welfare relative to the baseline in all four counterfactual scenarios. Note that all the counterfactual experiments alter the choice set available to prospective brides and grooms, either by changing the market competitiveness (measured by changes in sex ratios) or by changing the distribution of spousal characteristics (e.g., education or *hukou* status).

Figure 5: Welfare Comparisons of the Counterfactual Policy Experiments



Note: The figure plots changes in welfare for men and women by *hukou* status following each of the policy experiments described in Section 4.2. Data are from 2000 *Census* and 2005 *Mini-Census*.

Increasing Education for Rural Populations. This exercise explores the trade-off between educational achievement and social status. Despite the disadvantages of having a rural *hukou* status that is difficult to convert to an urban one, education is likely to have an impact on individuals' social division. According to Zhao (1997), education provides better access to higher-paying urban jobs for the rural population. Estimates from my baseline model also show

that high education is preferred by both men and women. To understand education's impact on the marriage market equilibrium, in this exercise, I first increase education for rural women. I also increase education for rural men as a comparison. Specifically, I set the education distribution of rural women (men) to be equal to that of urban women (men), which is in line with the goal of addressing the urban-rural gap in educational attainment underpinned by the *hukou* system. Figure 5 illustrates that equalizing the education level of rural women and urban women would result in a positive impact on the welfare of rural women and urban men. However, this would also reduce welfare for urban women and rural men. Conversely, if education were improved for rural men, the welfare change pattern would be reversed. Figure D11 shows that there is heterogeneity among rural men in terms of educational attainment: rural and below-high education men would lose the most when increasing education for rural women, while rural and high education men would benefit from the counterfactual policy. This is because rural and high-education men would now have a larger choice set with more high-education rural women. Also, because men dislike having spouses with higher education based on the model estimates, there would be less competition coming from urban and below-high education men. In contrast, high-education women, regardless of their *hukou* status, would suffer as a result of this counterfactual policy, since their competition would increase in the marriage market due to the increased supply of high-education women.

Changes in Sex Ratios. I also investigate the impact of changes in sex ratios under several scenarios. First, I increase the number of women in each province by 10% while keeping the number of men constant. Second, I reduce the number of women in each province by 10%, again keeping the number of men constant. Third, I apply the "natural" sex ratio of 105 males for every 100 females, following the WHO definition. I simulate the equilibrium marriage market outcomes for each scenario. When increasing the proportion of women, because urban women are still favored due to their urban status, the utility gain from having an urban status offsets the disadvantage from higher market competitiveness, as seen in Figure 5. Conversely, rural women experience a loss in welfare. The pattern is reversed when I reduce the proportion of women in the marriage market. A further decomposition by educational attainment, shown in Figure D12, suggests that there is significant heterogeneity in welfare effects when adjusting the sex ratios in the marriage market.

Removal of Status Differences. Ever since 2014, the Chinese government has endeavored to gradually eliminate the distinctions between urban and rural *hukou*. To replicate this goal, I perform a counterfactual analysis in which the status discrepancies have been totally abolished and predict the equilibrium outcomes of the marriage market. In this scenario, individuals only differ in the dimension of educational attainment. As depicted in Figure 5, this counterfactual policy is the only one that would yield welfare gains for all groups in the mar-

riage market. Furthermore, rural populations would experience a higher welfare gain compared to their urban counterparts. Figure D13 indicates that removing status differences would benefit all categories of individuals, even when the results are further decomposed by educational attainment. Nevertheless, one should cautiously interpret the welfare changes because the current analysis is confined to the marriage market and thus cannot measure changes in other domains, such as the labor market. The removal of the *hukou* barrier may also alter migration flows, which in turn would change the distribution of men and women in each marriage market. However, since the current model does not incorporate migration decisions, the estimates of welfare change presented in this study may be a lower bound. If migration flows were taken into account, individuals would have greater freedom in choosing their partners, resulting in a more significant change in welfare.

5 Conclusion and Discussions

This paper empirically investigates how social status shapes marital sorting by virtue of the urban-rural household registration (*hukou*) system in mainland China. Since the 1950s, this nationwide dichotomy of residential status has generated a far-reaching gap between urban and rural areas in many fundamental aspects of people's lives, including marriage. Building upon a 1998 policy reform that relaxed the restriction on children's status entitlement, I first use a difference-in-differences approach to demonstrate the weight of *hukou* status in marriage decisions. I then use a two-sided directed search and matching model to uncover the underlying mechanisms. The model allows me to capture the effect of the reform on market frictions, as well as to separately identify men's and women's preferences for status. Using the model estimates, I conduct a series of counterfactual policy experiments and quantify the welfare changes for each scenario.

My findings suggest that the *hukou* system indeed plays a significant role in marriage formation. Women have a stronger preference than men to match with a spouse who shares similar characteristics. Findings from subsequent counterfactual analysis demonstrate that if the Chinese government completely relaxed the *hukou* system, both men and women would benefit in the marriage market. Welfare gains would be the highest for women and for those in provinces with smaller fractions of urban populations.

Furthermore, it is interesting to note that while the 1998 reform led to a notable increase in intermarriages between urban men and rural women, and that such marriages are often seen as a means to improve social mobility for lower-class individuals, the welfare analysis conducted through a structural model indicates that both rural and urban women ended up worse off. In developing countries, where many women do not work outside their homes, marriage is arguably the single most important determinant of a woman's economic future. This paradoxical

outcome may be attributed to diminished bargaining power within households, and may inspire further research on the wider ramifications of unequal matches across individuals' social statuses.

This study has several limitations that open up possibilities for future research. First, I did not incorporate unobserved heterogeneity for preference parameters. The model assumes that individuals of different types have the same preference for a particular characteristic in the marriage decision. One extension would be to add unobserved heterogeneity in the deterministic utility model. However, adding unobserved heterogeneity could result in multiple equilibria issues in the matching game, creating identification challenges. Second, the model does not capture the impacts of social status on labor markets. It might likewise be intriguing to incorporate a migration model to address this topic. However, adding the migration stage would prevent marriage markets, now defined across provinces, from being independent of one another, which would complicate the current setting substantially. Moreover, estimating this new model would require more granular data, which is beyond the scope of this paper.

While further research could certainly complement the present findings, they still raise a range of immediate policy implications. One implication echoes the most recent *hukou* reform in China, implemented in 2014. The Chinese government has emphasized the urgency of eliminating the barriers induced by *hukou* on cross-regional population movement, especially in rural-urban migration. Meanwhile, it recently initiated policies to narrow down the discrepancies in the public welfare system between different *hukou* statuses. My paper offers new insights into the potential consequences of such policy reforms from the perspective of the marriage market. Once accounting for migration flows, the welfare gains from removing status differences would be even more substantial.

More broadly, this study provides a useful framework to study the impact of social status on marital outcomes and a new perspective on uncovering inequality in marriage markets. Specifically, the Chinese context may shed light on other Asian countries. For example, in India, where caste-based social hierarchies are rigidly enforced, policies that alter the inheritance rules for caste may be more effective in diminishing social segregation by reshaping the outcomes of marriage matching. This framework can also be extended to other scenarios, such as exploring how changes in immigration regulations or citizenship inheritance laws in Western countries, including the United States, may impact the sorting of spouses and affect educational and income inequalities between individuals from varying social backgrounds.

References

Abramitzky, R., Delavande, A., and Vasconcelos, L. (2011). Marrying up: The role of sex ratio in assortative matching. *American Economic Journal: Applied Economics*, 3(3):124–57.

Abramson, B. and Shayo, M. (2022). International integration and social identity. *Journal of International Economics*, 137:103577.

Adda, J., Pinotti, P., and Tura, G. (2020). There's more to marriage than love: the effect of legal status and cultural distance on intermarriages and separations. *CEPR Discussion Paper No. DP14432*.

Afridi, F., Li, S. X., and Ren, Y. (2015). Social identity and inequality: The impact of China's hukou system. *Journal of Public Economics*, 123:17–29.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Arcidiacono, P., Beauchamp, A., and McElroy, M. (2016). Terms of endearment: An equilibrium model of sex and matching. *Quantitative Economics*, 7(1):117–156.

Atkin, D., Colson-Sihra, E., and Shayo, M. (2021). How do we choose our identity? a revealed preference approach using food consumption. *Journal of Political Economy*, 129(4):1193–1251.

Banerjee, A., Duflo, E., Ghatak, M., and Lafortune, J. (2013). Marry for what? caste and mate selection in modern India. *American Economic Journal: Microeconomics*, 5(2):33–72.

Beauchamp, A., Calvi, R., and Fulford, S. (2021). Terms of engagement: Marriage and migration in India. *CEPR Discussion Paper No. DP16659*.

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. (1981). *A treatise on the family*. Cambridge, MA: Harvard University Press.

Bian, Y. (2002). Chinese social stratification and social mobility. *Annual Review of Sociology*, 28(1):91–116.

Brandt, L., Li, H., Turner, L., and Zou, J. (2018). Are China's "leftover women" really leftover? an investigation of marriage market penalties in modern-day China. Working paper, Department of Economics, University of Toronto.

Browning, M., Chiappori, P.-A., and Weiss, Y. (2014). *Economics of the Family*. Cambridge: Cambridge University Press.

Chade, H., Eeckhout, J., and Smith, L. (2017). Sorting through search and matching models in economics. *Journal of Economic Literature*, 55(2):493–544.

Chiappori, P.-A. (2017). *Matching with transfers: The economics of love and marriage*. Princeton University Press.

Chiappori, P.-A., Oreffice, S., and Quintana-Domeque, C. (2012). Fatter attraction: Anthropometric and socioeconomic matching on the marriage market. *Journal of Political Economy*, 120(4):659–695.

Chiappori, P.-A., Oreffice, S., and Quintana-Domeque, C. (2017). Bidimensional matching with heterogeneous preferences: Education and smoking in the marriage market. *Journal of the European Economic Association*, 16(1):161–198.

Chiappori, P.-A. and Salanié, B. (2016). The econometrics of matching models. *Journal of Economic Literature*, 54(3):832–61.

Chiappori, P.-A. and Salanié, B. (2021). Mating markets. *CEPR Discussion Paper No. DP16041*.

Choo, E. (2015). Dynamic marriage matching: An empirical framework. *Econometrica*, 83(4):1373–1423.

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1):175–201.

Cornelius, T. J. (2003). A search model of marriage and divorce. *Review of Economic Dynamics*, 6(1):135–155.

Devine, O. J., Louis, T. A., and Halloran, M. E. (1994). Empirical Bayes methods for stabilizing incidence rates before mapping. *Epidemiology*, 5(6):622–630.

Dong, B., Hsieh, Y.-W., and Zhang, X. (2020). Implementing maximum likelihood estimation of empirical matching models. *Computational Economics*, pages 1–32.

Dubé, J.-P., Fox, J. T., and Su, C.-L. (2012). Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica*, 80(5):2231–2267.

Dupuy, A. (2021). Migration in China: To work or to wed? *Journal of Applied Econometrics*, 36(4):393–415.

Dupuy, A. and Galichon, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, 122(6):1271–1319.

Ebenstein, A. (2011). Estimating a dynamic model of sex selection in China. *Demography*, 48(2):783.

Edlund, L. (1999). Son preference, sex ratios, and marriage patterns. *Journal of Political Economy*, 107(6):1275–1304.

Falk, A. and Hermle, J. (2018). Relationship of gender differences in preferences to economic development and gender equality. *Science*, 362(6412).

Fernandez, R., Guner, N., and Knowles, J. (2005). Love and money: A theoretical and empirical analysis of household sorting and inequality. *The Quarterly Journal of Economics*, 120(1):273–344.

Fershtman, C., Murphy, K. M., and Weiss, Y. (1996). Social status, education, and growth. *Journal of Political Economy*, 104(1):108–132.

Gong, P., Liang, S., Carlton, E. J., Jiang, Q., Wu, J., Wang, L., and Remais, J. V. (2012). Urbanisation and health in China. *The Lancet*, 379(9818):843–852.

Guilmoto, C. Z. (2011). Skewed Sex Ratios at Birth and Future Marriage Squeeze in China and India, 2005–2100. *Demography*, 49(1):77–100.

Han, L., Li, T., and Zhao, Y. (2015). How status inheritance rules affect marital sorting: Theory and evidence from urban China. *The Economic Journal*, 125(589):1850–1887.

Han, L. and Shi, X. (2019). How does intergenerational investment respond to changes in the marriage market? evidence from China. *Journal of Development Economics*, 139:109 – 121.

Hitsch, G. J., Hortaçsu, A., and Ariely, D. (2010a). Matching and sorting in online dating. *American Economic Review*, 100(1):130–63.

Hitsch, G. J., Hortaçsu, A., and Ariely, D. (2010b). What makes you click?—mate preferences in online dating. *Quantitative Marketing and Economics*, 8(4):393–427.

Hsieh, Y.-W. (2012). Understanding mate preferences from two-sided matching markets: Identification, estimation and policy analysis. Working paper, Department of Economics, University of Southern California.

Kalmijn, M. (1991). Status homogamy in the United States. *American Journal of Sociology*, 97(2):496–523.

Kane, T. J. and Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Technical report, National Bureau of Economic Research.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic Perspectives*, 34(3):174–92.

Li, H., Loyalka, P., Rozelle, S., and Wu, B. (2017). Human capital and China's future growth. *Journal of Economic Perspectives*, 31(1):25–48.

Lui, L. (2018). Marital power in inter-hukou families in China: An intersectionality approach. *Journal of Family Issues*, 39(5):1165–1190.

Meng, L. and Zhao, M. Q. (2019). Bride drain: An unintended consequence of China's urban-rural divide. *Labour Economics*, 58:69 – 80.

Meng, X. (2012). Labor market outcomes and reforms in China. *Journal of Economic Perspectives*, 26(4):75–102.

Meng, X. and Zhang, J. (2001). The two-tier labor market in urban China: Occupational segregation and wage differentials between urban residents and rural migrants in Shanghai. *Journal of Comparative Economics*, 29(3):485–504.

Nie, H. and Xing, C. (2011). When city boy falls in love with country girl: Baby's hukou, hukou reform, and inter-hukou marriage. In *Institute for the Study of Labor (IZA) annual workshop, Bonn, Germany*.

Oh, S. (2021). Does identity affect labor supply? *Available at SSRN 3998025*.

Ong, D. and Wang, J. (2015). Income attraction: An online dating field experiment. *Journal of Economic Behavior & Organization*, 111:13–22.

Ong, D., Yang, Y. A., and Zhang, J. (2020). Hard to get: The scarcity of women and the competition for high-income men in urban China. *Journal of Development Economics*, 144:102434.

Porter, M. (2016). How do sex ratios in China influence marriage decisions and intra-household resource allocation? *Review of Economics of the Household*, 14:337–371.

Qian, Y. and Qian, Z. (2017). Assortative mating by education and hukou in Shanghai. *Chinese Sociological Review*, 49(3):239–262.

Raiber, E., Ren, W., Bovet, J., Seabright, P., and Wang, C. (2023). What do parents want? parental spousal preferences in China. *Economic Development and Cultural Change*, Forthcoming.

Rozelle, S. and Boswell, M. (2021). Complicating China's rise: Rural underemployment. *The Washington Quarterly*, 44(2):61–74.

Sharygin, E., Ebenstein, A., and Das Gupta, M. (2013). Implications of China's future bride shortage for the geographical distribution and social protection needs of never-married men. *Population Studies*, 67(1):39–59.

Shimer, R. and Smith, L. (2000). Assortative matching and search. *Econometrica*, 68(2):343–369.

Song, Y. (2014). What should economists know about the current chinese hukou system? *China Economic Review*, 29:200 – 212.

Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–72.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press, 2nd edition.

Vandenberg, L. (2018). Finding love in china's marriage markets. *The Diplomat*. Available at: https://thediplomat.com/2018/11/finding-love-in-chinas-marriage-markets/ (Accessed: June 6th, 2020).

Vortherms, S. A. (2019). China's missing children: Political barriers to citizenship through the household registration system. *The China Quarterly*, 238:309–330.

Vortherms, S. A. (2021). Hukou as a case of multi-level citizenship. In *The Routledge Handbook of Chinese Citizenship*, pages 132–142. Routledge.

Wang, H., Dwyer-Lindgren, L., Lofgren, K. T., Rajaratnam, J. K., Marcus, J. R., Levin-Rector, A., Levitz, C. E., Lopez, A. D., and Murray, C. J. (2012). Age-specific and sex-specific mortality in 187 countries, 1970–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859):2071–2094.

Wang, P. (2021). Love in China (1950–now). In Mayer, C.-H. and Vanderheiden, E., editors, *International Handbook of Love: Transcultural and Transdisciplinary Perspectives*, pages 955–973. Springer International Publishing, Cham.

Wang, Y. and Schwartz, C. R. (2018). Hukou intermarriage and social exclusion in China. *Research in Social Stratification and Mobility*, 56:28–39.

Weiss, Y. (1997). Chapter 3 the formation and dissolution of families: Why marry? who marries whom? and what happens upon divorce. volume 1 of *Handbook of Population and Family Economics*, pages 81–123. Elsevier.

Weiss, Y. and Fershtman, C. (1998). Social status and economic performance:: A survey. *European Economic Review*, 42(3):801–820.

Wright, R., Kircher, P., Julien, B., and Guerrieri, V. (2021). Directed search and competitive search: A guided tour. *Journal of Economic Literature*, 59(1):90–148.

Wu, X. (2019). Inequality and social stratification in postsocialist China. *Annual Review of Sociology*, 45(1):363–382.

Wu, X. and Treiman, D. J. (2004). The household registration system and social stratification in China: 1955–1996. *Demography*, 41(2):363–384.

Xiang, J. (2015). Gendered pathways to hukou status mobility in China. *Chinese Journal of Sociology*, 1(2):231–253.

Yu, J. and Xie, Y. (2015). Changes in the determinants of marriage entry in post-reform urban China. *Demography*, 52(6):1869–1892.

Zhang, L., Huang, J., and Rozelle, S. (2002). Employment, emerging labor markets, and the role of education in rural China. *China Economic Review*, 13(2):313–328.

Zhang, X. and Kanbur, R. (2005). Spatial inequality in education and health care in China. *China Economic Review*, 16(2):189–204.

Zhao, Y. (1997). Labor migration and returns to rural education in China. *American Journal of Agricultural Economics*, 79(4):1278–1287.

Zhu, J., Cui, L., Wang, K., Xie, C., Sun, N., Xu, F., Tang, Q., and Sun, C. (2019). Mortality pattern trends and disparities among Chinese from 2004 to 2016. *BMC Public Health*, 19:1–7.

ONLINE APPENDIX

Appendix A. Calculate Relevant Sex Ratio

Since the sex ratio of each province is a macro-level measure, I rely on the 2000 *Census* from the National Bureau of Statistics of China to calculate the sex ratios for more granular marriage markets defined by *hukou* status, educational group, and province. The 2000 *Census* is also the most recent Population Census available that can be used to predict marriage outcomes in 2005. To deal with extremely unbalanced sex ratios in a certain cell defined by the specific types, I apply the Bayesian shrinkage method to alleviate the concern of imbalanced distribution of sex ratio (Devine et al., 1994; Kane and Staiger, 2008). It is a method that has been commonly used in the literature on estimating teacher value-added. The empirical Bayes estimator is a weighted combination of the observed rate and an expected rate, usually the estimated mean of the prior distribution. It involves two steps: First, to estimate the overall distribution of the data; Second, to use that distribution as the prior for estimating each average.

In my exercise, I apply three different approaches corresponding to three different prior distributions. The first approach is the Beta-Binomial approach. It relies on the fact that the beta distribution is the conjugate prior to the binomial. I assume that for each demographic group $d = \{p, g, h\}$, which is the combination of province p, gender g, *hukou* type h, the probability of having high education, i.e., at least middle school, is denoted as p_d. For each demographic group d, the observed rate of high education:

$$p_d = \frac{\text{\# actually holding at least middle school degree}}{\text{total \# of observations}} = \frac{s_d}{s_d + f_d},$$

where s_d is the observed actual people holding at least a middle school degree and f_d is the observed actual observations holding degrees lower than middle school. The beta-binomial approach would adjust this number to

$$\tilde{p}_d = \frac{s_d + s_0}{s_d + f_d + (s_0 + f_0)}, \tag{12}$$

where s_0 and f_0 are the two parameters of the β distribution, which we can think of as "banked" successes and failures.

For implementation, I follow two steps. First, I use population data to get the prior. Following the Bayes theorem, if we used a β distribution with parameters (s_0, f_0) as a prior, then the a posteriori distribution is also a β distribution with parameters $(s_0 + s_d, f_0 + f_d)$. To find estimates of the two parameters s_0 and f_0 of the β distribution, one way of doing this is to use the mean μ and variance σ^2 of the distribution of the true data, given by:

$$s_0 = \mu \left(\frac{\mu(1-\mu)}{\sigma^2} - 1 \right), \quad f_0 = (1-\mu) \left(\frac{\mu(1-\mu)}{\sigma^2} - 1 \right).$$

The second step is to use prior to "shrink" estimates to population values. Once equipped with the estimates of s_0 and f_0, I then follow (12) to "shrink" the raw proportions in demographic group d to the expected value.

Since this approach starts with a relatively strong assumption that all rates of high educational attainment for the demographic groups are drawn from a single distribution, I next try an alternative prior distribution assumption that allows for more heterogeneity. In the second approach, i.e., the Bayesian Hierarchical approach, I assume priors for each demographic group are not fixed, but rather depend on other latent variables. I specify the following structure on μ but keep σ^2 constant

$$\log(\frac{\mu^{(d)}}{1-\mu^{(d)}}) = \lambda_{\text{district}} + \lambda_{\text{gender}} + \lambda_{\text{hukou}} + \lambda_{\text{gender}*\text{hukou}}.$$

The third approach I tested is the Dirichlet-Multinomial approach. Dirichlet is the conjugate prior to the multinomial distribution. This approach naturally extends the Beta-Binomial conjugation to the multiple discrete cases and posterior distribution follows a similar intuition. I take the simplest approach that categorizes 4 possible outcomes with gender×education as shown in Table 5.

Table 5: The Dirichlet-multinomial Distribution

Women, Low educ	Women, High educ
Men, Low educ	Men, High educ

After comparing the goodness of fit results for the three different prior distributions, I find that Dirichlet-Multinomial prior leads to the smallest prediction error, as shown in Figure D14.

Appendix B. Robustness Checks of the DID Analysis

In this section, I explore alternative specifications of the DID Analysis in Section 2.3.1 as a complement and robustness check. First, I test the effects based on the specification motivated by Nie and Xing (2011):

$$Y_{ic} = \beta_1 \text{post} + \beta_2 \text{post} \times G_c + \mathbf{X}_{ic}\boldsymbol{\mu} + \lambda_c + \varepsilon_{ic}, \tag{13}$$

where Y_{ic} denotes the outcome of a couple i in province c, $post = 1$ if i got married after 1998,[36] G_c is the absolute value of the difference between the urban and rural population in province c in the year 1998, as defined in the main text, $\mathbf{X}_{ic}$ is a vector of covariates including individual characteristics, such as educational attainment, marriage age, and whether belongs to a minority group. I also include province fixed effects λ_c. I estimate (13) using the same sample of married couples and explore the same marital outcomes as in Section 2.3.1. Table D4 shows the results of the impact of the policy on *hukou* intermarriages. It is evident that the results display the same pattern as those obtained using the main specification: the 1998 policy increased *hukou* intermarriages, with a more prominent effect on intermarriages between rural women and urban men.

Second, to mitigate concerns about the possible influence of the reform on part of the control group as defined in (1), I reconstructed both the treatment and control groups. The treatment

[36] Again, I view this exercise as suggestive evidence given that there is the potential for endogenous timing of marriage.

group consisted of men younger than 22 years in 1998, and those older than 22 but single at that time. The remaining sample constituted the control group. The results, presented in Table D5, remain robust to this additional robustness check.

Third, to test the sensitivity of the sample selection concern, I also try estimating the main specification (1) by constructing a larger sample including all couples in the households. I match husbands and wives based on the reported relationships with the household head[37] and the time of marriage (year and month). I also rely on gender to identify whether the person is the husband or the wife. After imposing the same constraints as in the main analysis,[38] the total number of couples in the final analysis sample is 136,747, including 101,630 cases of household head with the spouse, 38 cases of father with mother (or father-in-law and mother-in-law), and 35,079 cases of son and daughter-in-law (or daughter and son-in-law). Results in Table D6 show the same pattern.

Appendix C. More on Identification

Regarding the parameters of the deterministic utility function α, it is identified through the variation of sex ratios in each marriage market defined by the cross-product of characteristics. As there are in sum $6 \times 6 \times 28$ marriage markets, the matches based on different characteristics can provide enough variations to pin down the preference parameters. To understand the identification of α, as one example, I plot the density of the same educational attainment couples, i.e., those that both husband and wife belong to the same educational group, among all the marriages for various of α_1 as shown in Figure D15(a). The results are based on 100 simulations by varying only α_1 and keeping all the other parameters fixed. We can see that when the preference parameter increases from a negative value to some positive value, the proportion of same educational attainment couples increases. This trend is consistent with model predictions. The identification of the other preference parameters follows the same argument.

To identify ρ, the elasticity in the production function, I rely on the distribution of aggregate matches across the markets. The variation in matches across provinces with different sex ratios contributes to the identification. Figure D15(b) plots the density of aggregate marriage rates when varying the value of elasticity of substitution parameter ρ. When the absolute value of ρ becomes smaller, the aggregate match rates decrease. This pattern provides a source of identification for ρ. For identification of parameters in the search friction, the key parameter in the model, technology change due to the policy for rural men relative to urban men, A_3, is identified following the exogenous variation of the 1998 policy change. The average level of matching technology, A_0, is identified based on the variation in matches across provinces with different sex ratios and the overall match rate in multiple provinces.[39] The technology change due to the policy for all groups, A_1, is identified based on the variation in the age in the estimation sample. Similarly, for the pre-policy technology difference for rural men compared

[37] This information includes household head himself (herself), spouse, son or daughter, father or mother, father-in-law or mother-in-law, grandfather or grandmother, son-in-law or daughter-in-law, grandchild, brother or sister, and others.

[38] These include (i) focusing on current marriages, (ii) restricting to those who are married at or above the legal marriage age in China, and (iii) restricting to couples who married after 1990.

[39] An important note is that identification of search friction parameter A_0 requires the inclusion of multiple provinces. If the third dimension of the marriage market, defined by the number of provinces in the empirical context, is too small, it will fail to identify A_0 due to not-enough variations.

to other groups, A_2 is identified based on the variation in rural and urban status in the sample. Following the traditional practice of the standard discrete choice logit model, I normalize σ to be equal to one in estimation.

Appendix D. Supplementary Tables and Figures

Table D1: The Matching Distribution by Type of Women

	Same type	Marry up (educ)	Marry down (educ)	Marry up (hukou)	Marry down (hukou)
All married women	.48	.11	.30	.04	.02
Hukou Educ					
Rural Low	.35	.01	.58	.03	.00
Rural Mid	.59	.14	.13	.09	.00
Rural High	.13	.47	.14	.46	.00
Urban Low	.18	.01	.71	.00	.16
Urban Mid	.48	.11	.35	.00	.08
Urban High	.51	.29	.19	.00	.02

Note: This table reports the matching pattern by type of women. The Type is defined as the combination of *hukou* status, either rural or urban, and the three educational groups: low, middle, or high. Education low: primary school or below; mid: above primary but below high school or high school; high: above high school. Marry-up" ("Marry-down") refers to the case when the matched spouse has more (less) desirable traits than the individual. Table entries are cell percentages. *Data Source:* National representative data from 2005 *Mini-Census*.

Table D2: Identification

Parameter	Model fundament	Data source
A_0	The average level of matching technology	Variation in matches across provinces with different sex ratios and the overall match rate in multiple provinces
A_1	Technology change due to the policy for all groups	Dummy for those younger than legal marriage age, 22
A_2	Pre-policy technology difference for rural men compared to other groups	Dummy for rural men
A_3	Technology change due to the policy for rural men	Interaction of rural men by the "treatment" group
ρ	Elasticity in the production function	Variation in matches across provinces with different sex ratios and the overall match rate
α	Preference for a specific characteristic	Variation of sex ratios on each market defined by the cross product of characteristics

Note: This table summarizes the identification strategy of the parameters in the model.

Table D3: Correlation of Welfare Change with Province Characteristics

	Winner proportion$_m$
Urban pop/rural pop	-0.015***
	(0.004)
Central region	0.003
	(0.005)
Western region	0.029***
	(0.009)
log(total population)	-0.012***
	(0.004)
$\lvert Urban_{pop} - Rural_{pop} \rvert$	0.043***
	(0.014)
Mean Dep. Var.	0.029
R^2	0.744
N	28

Note: This table shows the regression results of the proportion of male winners in each province of China on the ratio of urban-to-rural population and several province-level characteristics, including indicators of location in the central or western region, the logarithm of the province's total population, and the absolute value of the gap in population share between urban and rural areas. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$. *Data Source:* 2000 *Census* and 2005 *Mini-Census*.

Table D4: The Impact of *Hukou* Policy on Intermarriage: Alternative Specification

	(1) intermarriage	(2) intermarriage	(3) rural w -urban m	(4) rural w-urban m	(5) rural m-urban w	(6) rural m-urban w
Post	0.122***	0.114***	0.079***	0.077***	0.043***	0.036***
	(0.011)	(0.011)	(0.009)	(0.009)	(0.006)	(0.006)
$Post \times G_c$	-0.021*	-0.027**	-0.015*	-0.018*	-0.006	-0.009
	(0.012)	(0.012)	(0.009)	(0.010)	(0.006)	(0.006)
Individual controls	No	Yes	No	Yes	No	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
p-value for (3)=(5): 0.427						
p-value for (4)=(6): 0.398						
Mean Dep. Var.	0.065	0.065	0.040	0.040	0.024	0.024
R^2	0.016	0.020	0.011	0.013	0.005	0.008
N	82,679	82,679	82,679	82,679	82,679	82,679

Note: This table shows the results of estimating Eq(13). Intermarriage is defined as marriages between urban and rural *hukou*. G_c is the absolute value of the difference between urban and rural population share in province c in the year 1998. $post = 1$ if the couple i got married after 1998. Individual controls include the woman's marriage age, education level, and whether belongs to a minority group. I restrict samples to those who got married after 1990. I also test the equality of β_2 for both types of intermarriages and report the corresponding p-values. Robust standard errors, clustered at the city level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$. *Data Source:* 2005 *Mini-Census*.

Table D5: The Impact of *Hukou* Policy on Intermarriage: Robustness Check

	(1) intermarriage	(2) intermarriage	(3) rural w -urban m	(4) rural w-urban m	(5) rural m-urban w	(6) rural m-urban w
$\mathbb{1}(Age^{1998} < 22)$	0.123***	0.113***	0.080***	0.080***	0.043***	0.033***
	(0.010)	(0.010)	(0.008)	(0.008)	(0.005)	(0.005)
$\mathbb{1}(Age^{1998} < 22) \times G_c$	-0.032***	-0.032***	-0.024***	-0.030***	-0.007*	-0.002
	(0.007)	(0.007)	(0.006)	(0.006)	(0.004)	(0.004)
Individual controls	No	Yes	No	Yes	No	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
p-value for (3)=(5): 0.017						
p-value for (4)=(6): 0.000						
Mean Dep. Var.	0.065	0.065	0.040	0.040	0.024	0.024
R^2	0.016	0.020	0.011	0.014	0.005	0.008
N	82,679	82,679	82,679	82,679	82,679	82,679

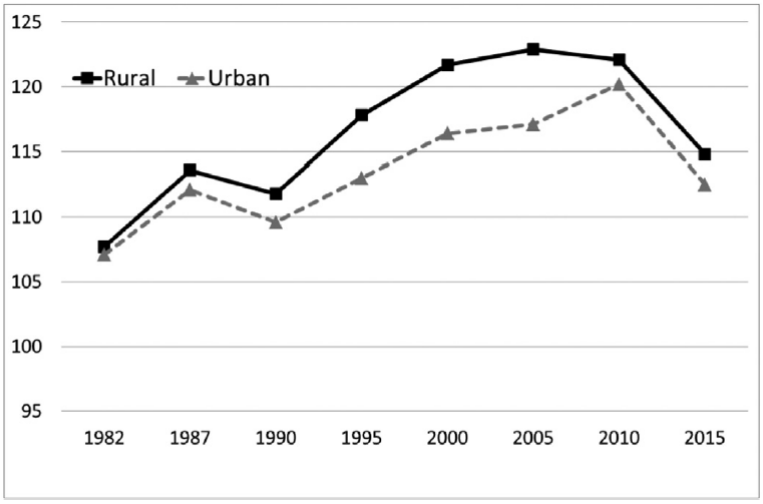
Note: This table shows the robustness check results of the DID analysis of Eq(1) by reconstructing the treatment and control groups. In this analysis, the treatment group was defined as comprising men who were younger than 22 years of age in 1998, as well as those who were older than 22 in that year but were single at the time. The control group was defined as comprising the remaining sample. Intermarriage is defined as marriages between urban and rural *hukou*. G_c is the absolute value of the difference between urban and rural population share in province c in the year 1998. Individual controls include the woman's marriage age, education level, and whether belongs to a minority group. I restrict samples to those who got married after 1990. I also test the equality of β_2 for both types of intermarriages and report the corresponding p-values. Robust standard errors, clustered at the city level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$. *Data Source:* 2005 *Mini-Census*.

Table D6: The Impact of *Hukou* Policy on Intermarriage: Alternative Sample

	(1) intermarriage	(2) intermarriage	(3) rural w -urban m	(4) rural w-urban m	(5) rural m-urban w	(6) rural m-urban w
$\mathbb{1}(Age^{1998} < 22)$	0.014*	0.030***	0.009	0.025***	0.006	0.005
	(0.007)	(0.008)	(0.006)	(0.006)	(0.004)	(0.004)
$\mathbb{1}(Age^{1998} < 22) \times G_c$	-0.069***	-0.075***	-0.048***	-0.053***	-0.021***	-0.021***
	(0.013)	(0.013)	(0.010)	(0.010)	(0.007)	(0.007)
Individual controls	No	Yes	No	Yes	No	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
p-value for (3)=(5): 0.011						
p-value for (4)=(6): 0.003						
Mean Dep. Var.	0.068	0.068	0.045	0.045	0.024	0.024
R^2	0.016	0.020	0.012	0.018	0.005	0.005
N	136,747	136,747	136,747	136,747	136,747	136,747

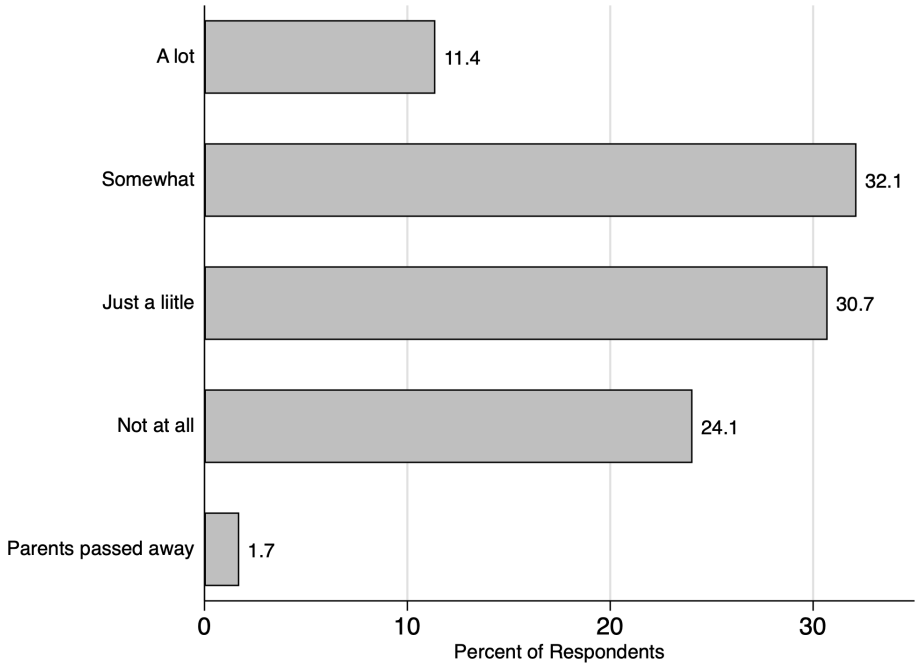
Note: This table shows the results of estimating Eq(1) using the full sample of married couples. Intermarriage is defined as marriages between urban and rural *hukou*. G_c is the absolute value of the difference between urban and rural population share in province c in 1998. Individual controls include the man's marriage age, education level, and whether belongs to a minority group. I restrict samples to those who got married after 1990 for all the specifications. I also test the equality of β_2 for both types of intermarriages and report the corresponding p-values. Robust standard errors, clustered at the city level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$. The data is from 2005 *Mini-Census*.

Figure D1: Sex ratio at birth, by urban-rural, 1982–2015



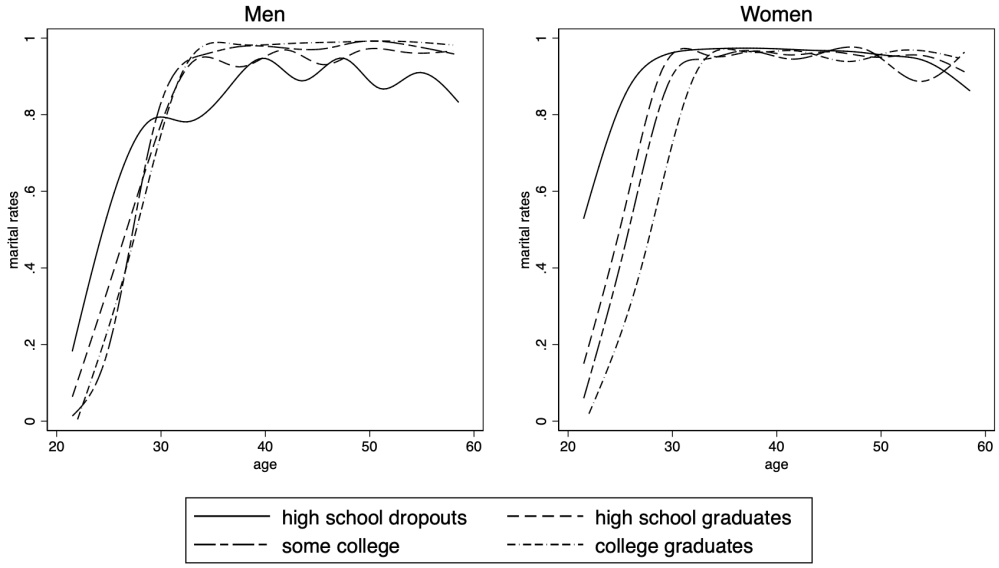
Note: Figure 1 in Meng and Zhao (2019).

Figure D2: The Role of Parents in Marriage Decisions



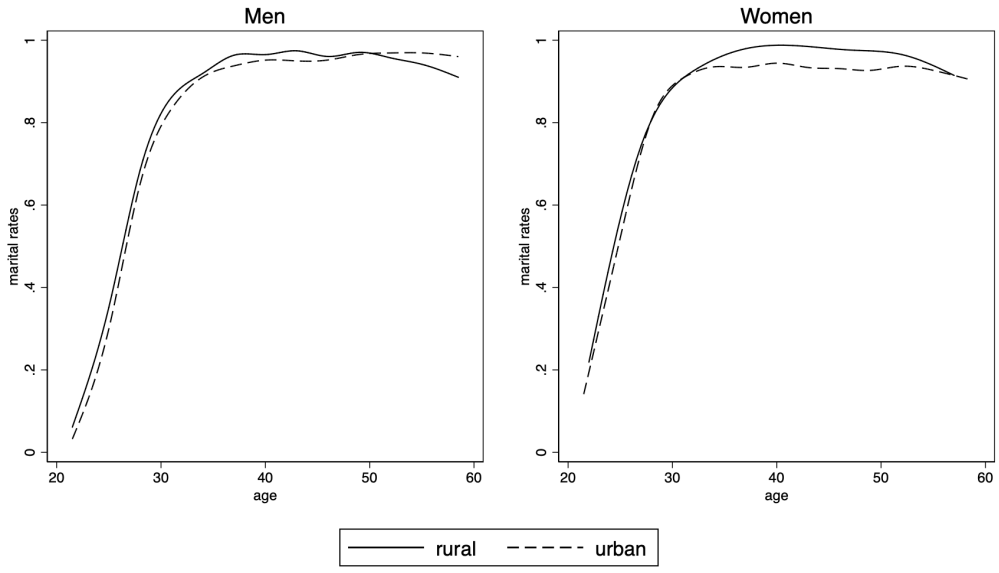
Note: This figure visualizes the answers to the survey question “To what extent did your parents’ opinion affect your marriage decisions?" Data are from the 2017 China General Social Survey.

Figure D3: Marital Rates by Educational Attainment and Gender



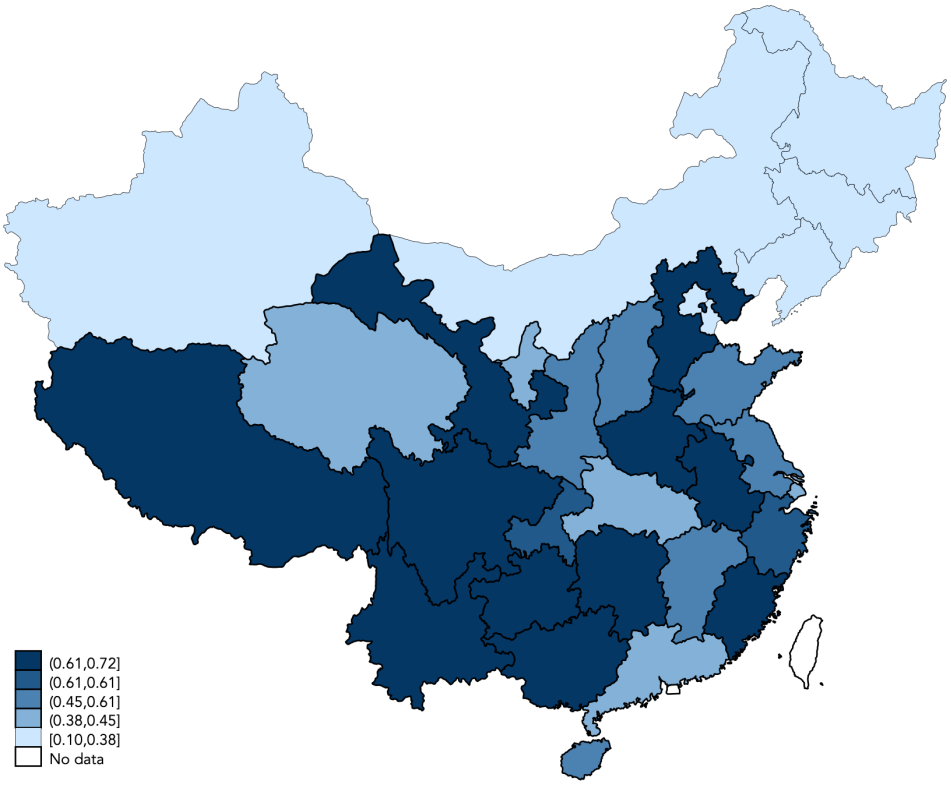
Note: This figure shows the marital rates by educational attainment and gender. Data are from 2005 *Mini-Census*.

Figure D4: Marital Rates by Hukou Type and Gender



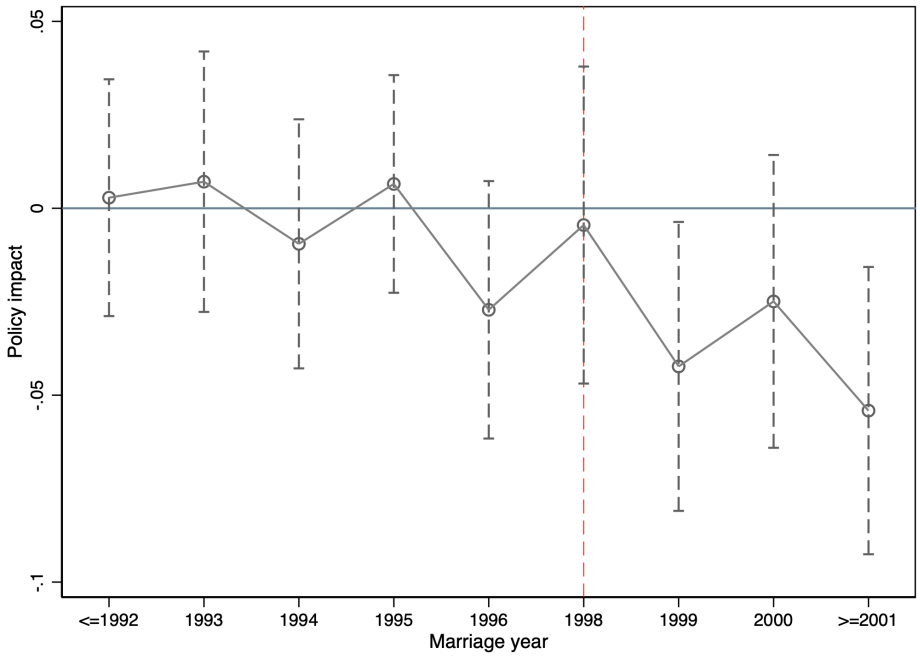
Note: This figure shows the marital rates by *hukou* type and gender. Data are from 2005 *Mini-Census*.

Figure D5: The Gap between Urban and Rural Populations by Province



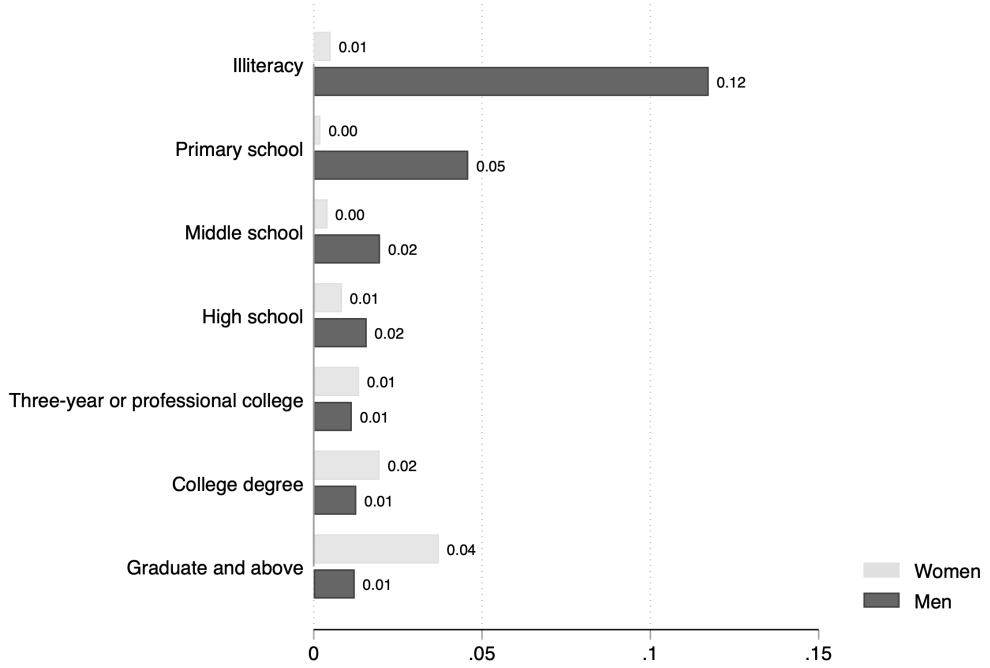
Note: This figure shows the variation in terms of the urban and rural population "gap" in each province in China. Darker color means a more imbalanced composition of the urban and rural populations. Data are from the 1998 China Population and Employment Statistics Yearbook.

Figure D6: Validation of the Parallel Trend Assumption



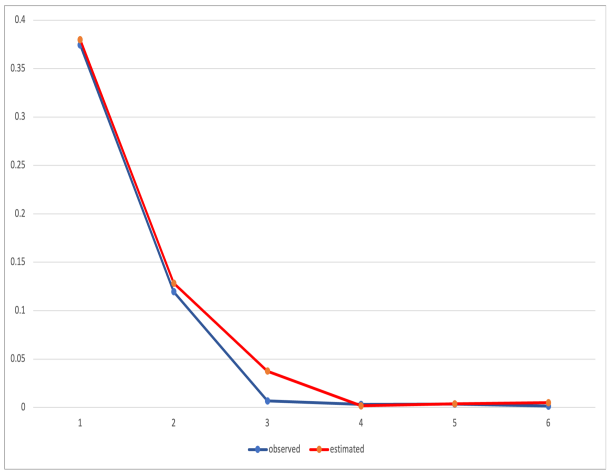
Note: This figure visualizes the impact of the policy on intermarriage rates by specific marriage year to test the parallel trend assumption of the difference-in-differences (DID) analysis presented in Eq(1). Intermarriage is defined as marriages between urban and rural *hukou*. Data are from 2000 *Census* and 2005 *Mini-Census*.

Figure D7: Proportion of Singles by Education Level and Gender

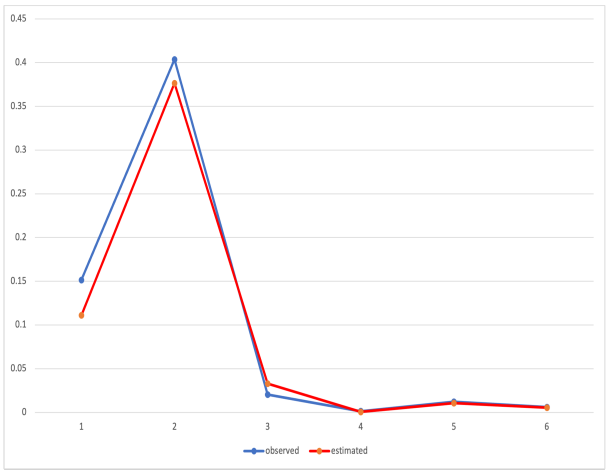


Note: Age group: 35+ years old. The numerator is the total number of unmarried men (women) that belong to a particular educational level group, while the denominator represents the total number of men (women) in that same educational attainment level group. Data are from 2005 *Mini-Census*.

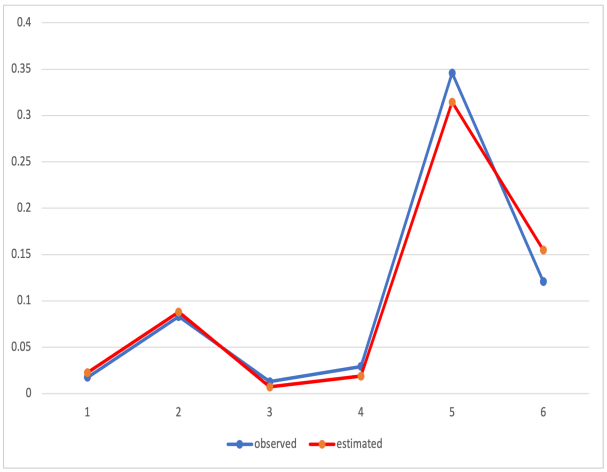
Figure D8: Model Goodness of Fit



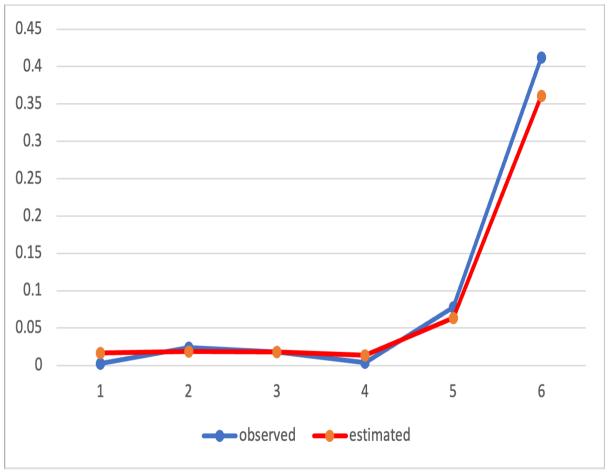
(a) Match rates for rural and low education men



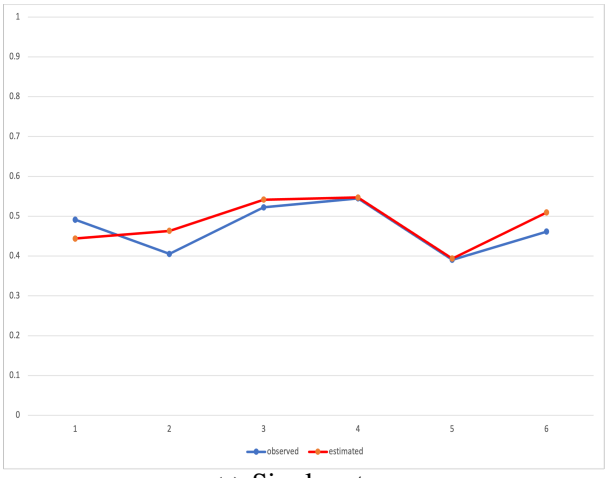
(b) Match rates for rural and middle education men



(c) Match rates for urban and middle education men



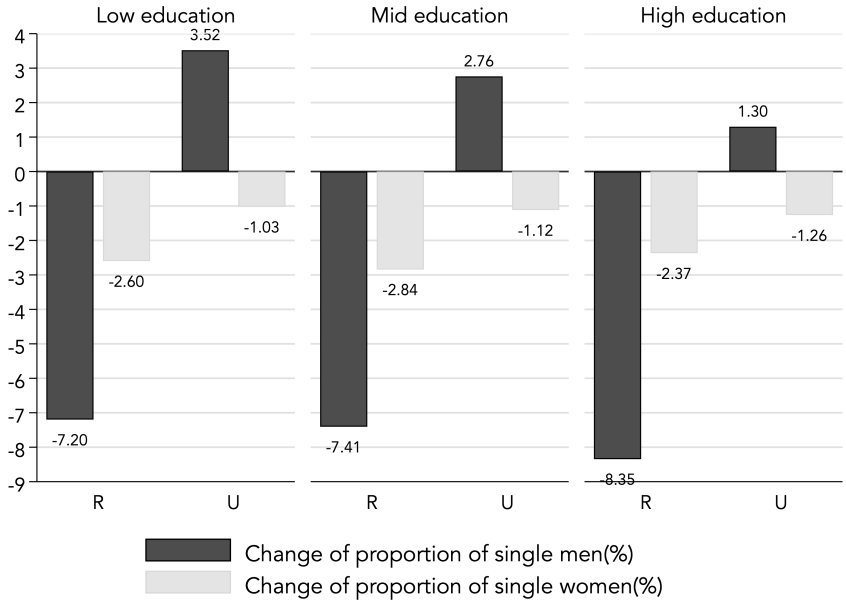
(d) Match rates for urban and high education men



(e) Single rates

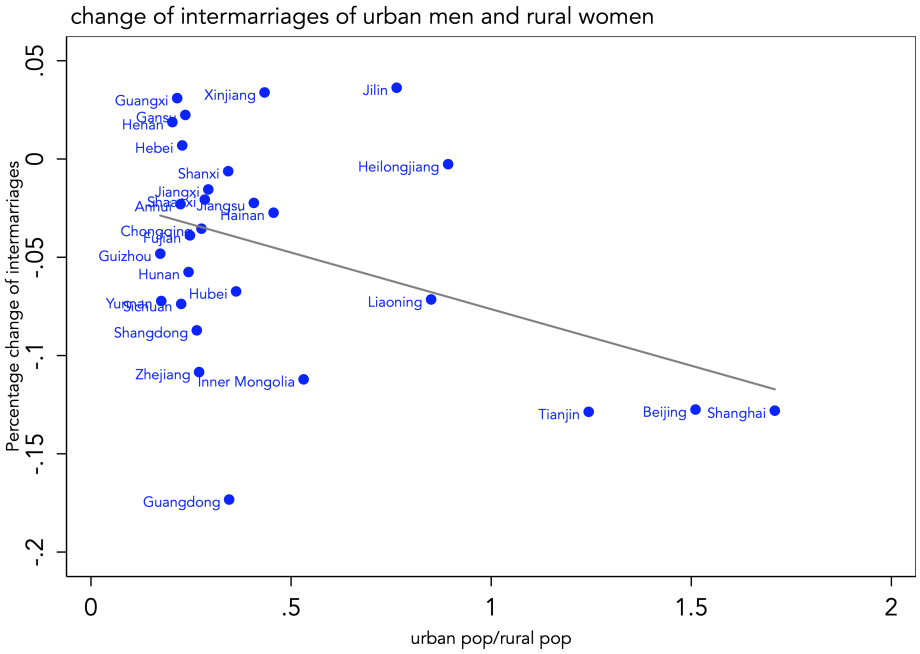
Note: These figures show the goodness of fit of the estimated model. The X-axis denotes the specific type of individual. The Y-axis denotes the matched rates. For presentation purposes, I only show the goodness of fit plots for four types of individuals and the single rates. The model fits the matched rates for the other two types of individuals, i.e., rural and high education men and urban and low education men well. The simulated results are obtained based on 100 simulations. Data are from 2000 *Census* and 2005 *Mini-Census*.

Figure D9: The Welfare Effect of the 1998 Policy Absence as Measured by Changes in Single Rates



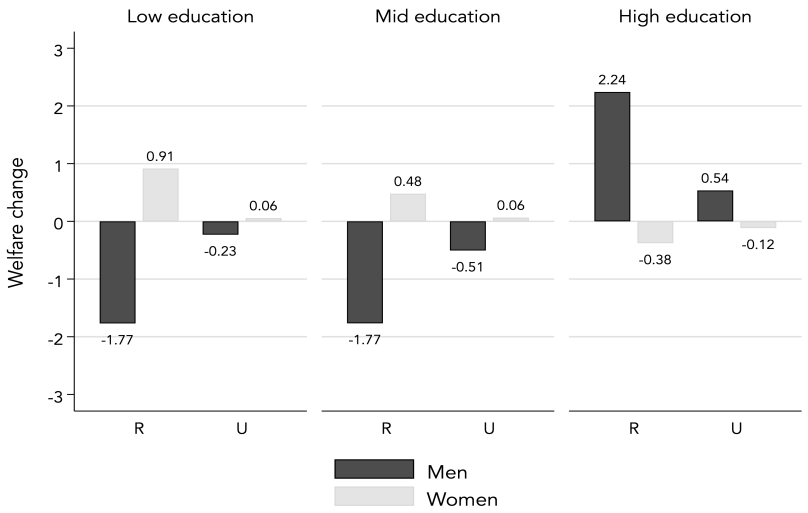
Note: This figure shows the changes in single rates for different types of individuals by educational attainment and *hukou* status in the absence of the 1998 policy. Data are from 2000 *Census* and 2005 *Mini-Census*.

Figure D10: Change of Intermarriages in the Absence of 1998 Policy by Province

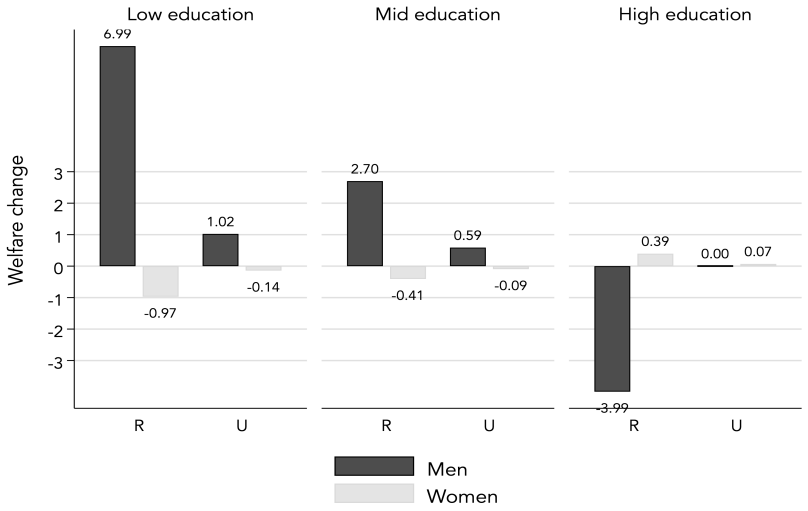


Note: This figure shows the change in intermarriage rates between urban men and rural women at the province level. The X-axis represents the ratio of urban *hukou* populations to rural *hukou* populations. Data are from 2000 *Census* and 2005 *Mini-Census*.

Figure D11: Welfare Change When Increasing Education for Rural Populations



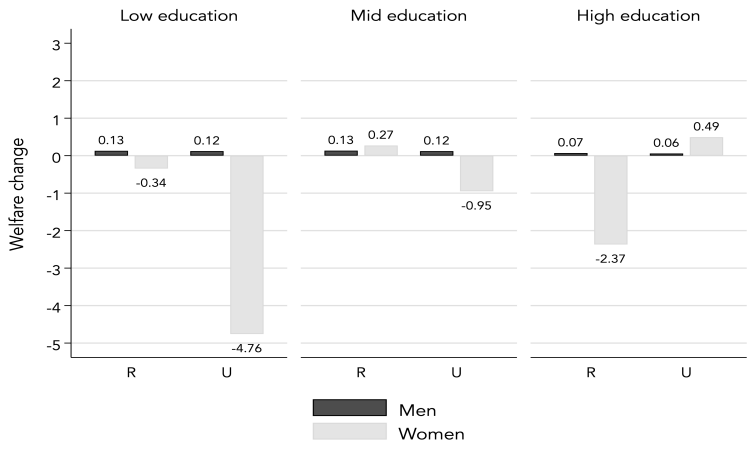
(a) Increase education for rural women



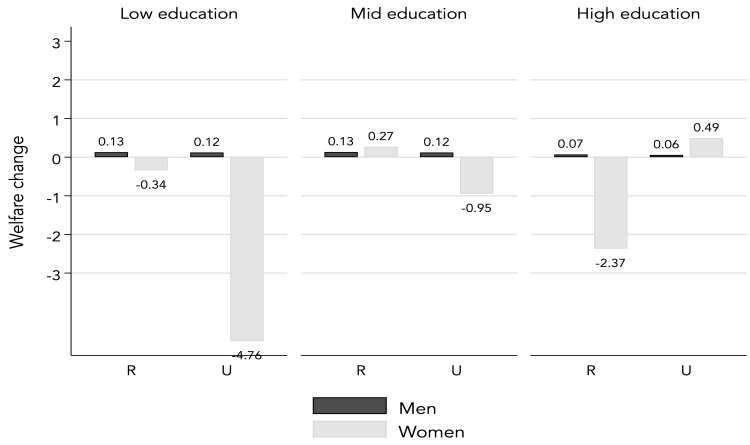
(b) Increase education for rural men

Note: This figure shows the welfare change for different types of individuals by educational attainment and *hukou* status for the counterfactual scenario when increasing education for rural populations. Data are from 2000 *Census* and 2005 *Mini-Census*.

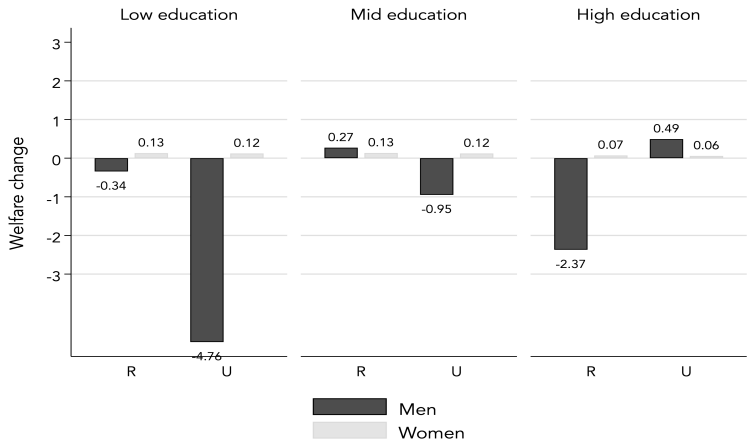
Figure D12: Welfare Change When Changing the Sex Ratios



(a) Increase proportion of women



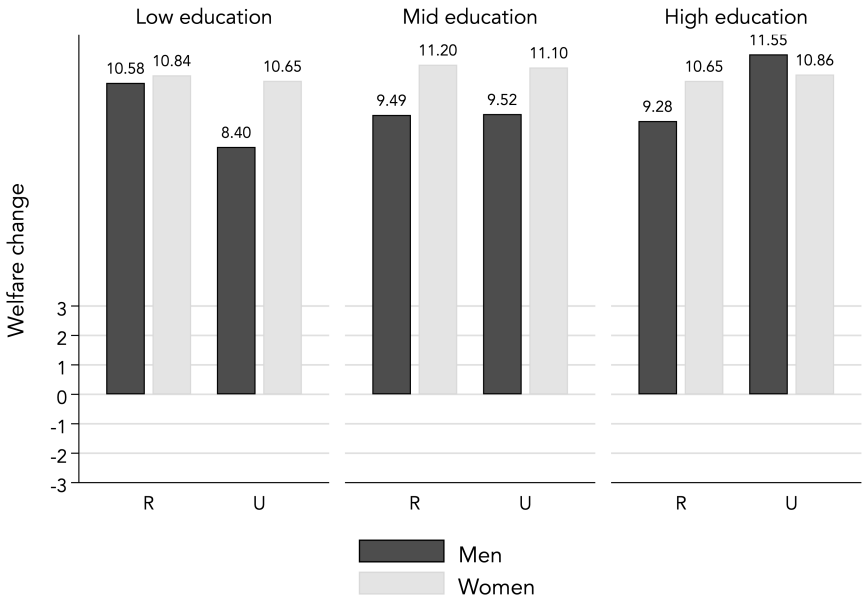
(b) Decrease proportion of women



(c) "Natural" sex ratio

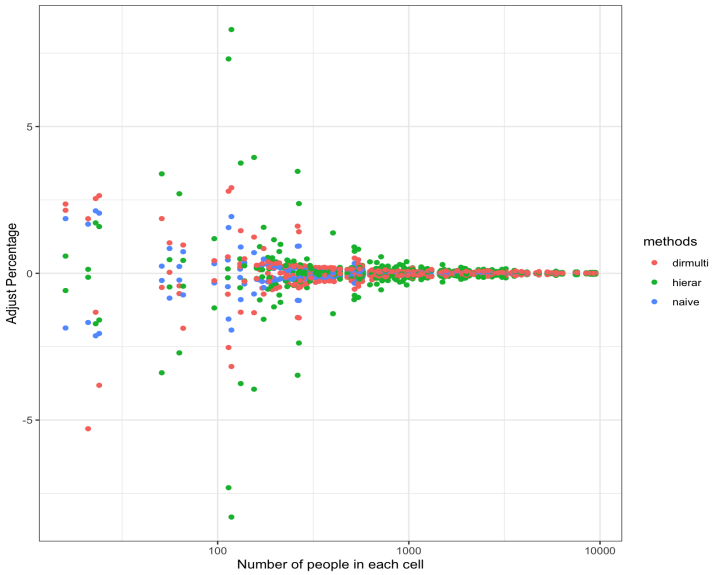
Note: This figure shows the welfare change for different types of individuals by educational attainment and *hukou* status when changing sex ratios. Data are from 2000 *Census* and 2005 *Mini-Census*.

Figure D13: Welfare Change When No *Hukou* Differences

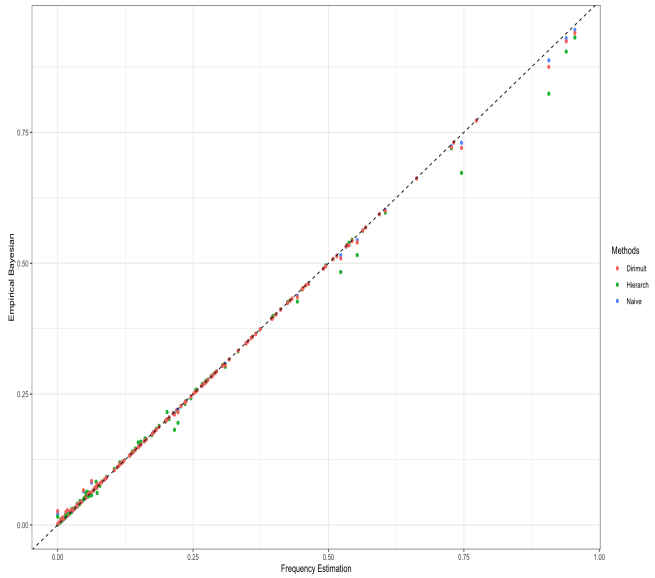


Note: This figure shows the welfare change for different types of individuals by educational attainment and *hukou* status when completely removing *hukou* differences. Data are from 2000 *Census* and 2005 *Mini-Census*.

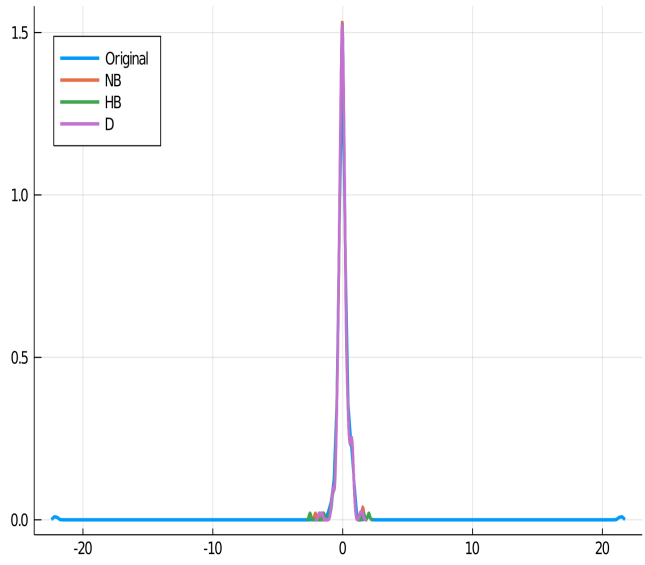
Figure D14: Graph Illustration of Empirical Bayesian Approach



(a) Comparison of the shrinkage adjustment



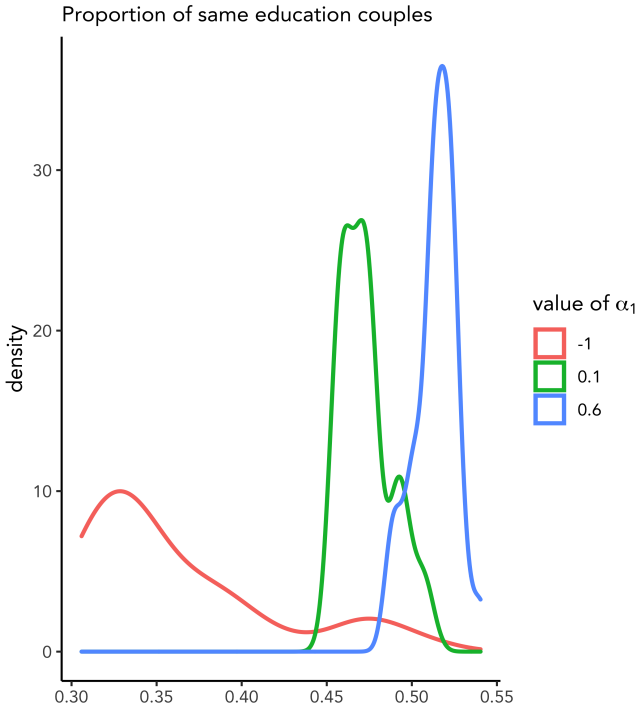
(b) Compare with frequency estimation



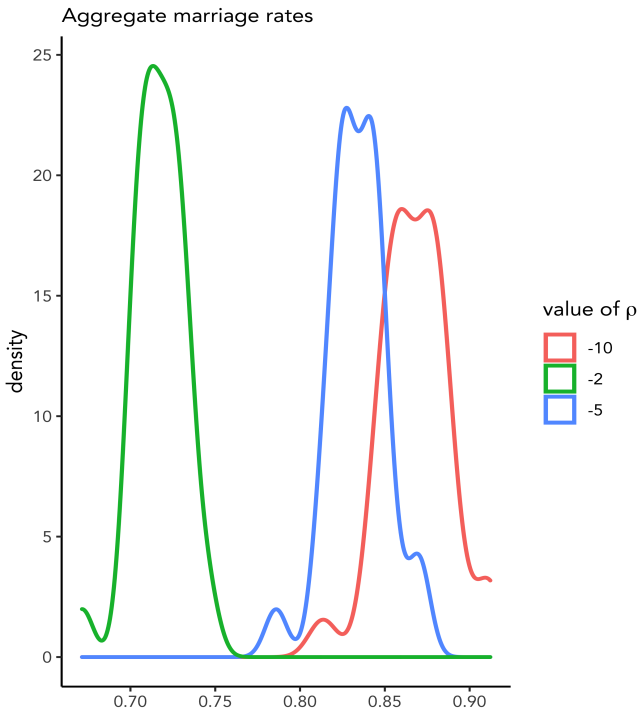
(c) Density of log(sex ratio)

Note: These figures visualize the results of the empirical Bayesian smoothing approach. Data are from 2000 *Census* and 2005 *Mini-Census*.

Figure D15: Graph Illustration of Identification of α and ρ



(a) Identification of α_1



(b) Identification of ρ

Note: These are the density plots of corresponding moments associated with changing values of a specific parameter while holding others constant. I provide examples of identifying two parameters, α and ρ. Results are obtained based on 100 simulations. Data are from 2000 *Census* and 2005 *Mini-Census*.