Hard to get:
The scarcity of women and the competition for high-income men in urban China[*]

November 16, 2018

David Ong, Yu (Alan) Yang, and Junsen Zhang[**]

Reports of the difficulties of elite women in finding suitable mates have been increasing despite the growing scarcity of women in China. We show that this phenomenon can be a consequence of women's preference for men who have higher incomes than themselves. With such a reference-dependent preference, the pool of preferred men shrinks as women's income increases, while the pool of competing poorer women expands. For high-income (h-)women, even when high-income (H-)men are more plentiful and richer (as in China), the direct effect of a greater number of desirable men can be overwhelmed by the indirect effect of the competitive "entry" of low-income (l-)women. We test for these competitive effects using online dating field experimental, Census, and China Family Panel Studies data. Consistent with competitive entry, the search intensity of beautiful l-women for H-men increases with the sex ratio and the income of H-men, while that of the plain-looking decreases. Notwithstanding the search intensity of h-women for H-men uniformly increases, their probability of marriage decreases with the sex ratio and men's income relative to low-income women. The decrease is absolute for of plain-looking high-income women. Our evidence is consistent with intra-female competition for spouses who can cover the labour market opportunity cost of marriage and childbirth, which vary with women's income.

JEL Codes: C93, J01, J12

Keywords: marriage, sex ratio, online dating, reference-dependent preference, competition, beauty

---

[*] We thank numerous scholars and participants at the Society for the Economics of Households 2018, the Conference in Honor of Mark Rosenzweig 2017, Advances in Field Experiments Conference 2017, the Asian and Australasian Society of Labour Economics Conference 2017, Tokyo Labor Economics Conference 2016, the NBER China Economy Working Group Conference 2016, the Chinese University of Hong Kong Economics Department Workshop on Family and Labor Economics 2016, the Asian Meeting of the Econometrics Society 2016, the Chinese Economic Association Meeting 2016, and seminar participants at the economics department of the Shanghai University of Finance and Economics, Harbin Institute of Technology, McMaster University, Wilfrid Laurier University, Hong Kong University, and Lingnan University for their helpful comments.

[**] Junsen Zhang (corresponding author), jszhang@cuhk.edu.hk, Department of Economics, Chinese University of Hong Kong, Shatin, NT, Hong Kong

I. Introduction

Reports in the popular press (Fincher, 2012) and in the academic literature (Qian & Qian, 2014) of the difficulties of elite women in finding suitable mates have been increasing, despite the growing scarcity of marriageable women in China (Jiang, Feldman, & Li, 2014). This scarcity is partly the consequence of one of the most radical family planning experiments in history. Initiated in 1979, the one-child policy has resulted in hundreds of million fewer births in China. Owing to the traditional Chinese son preference, this decrease in births has not been equally distributed; at least 30 million women are now missing from the prime-age marriage market.

One might have supposed that the surviving women can only benefit from their scarcity. Indeed, this outcome is predicted by established economic theory; the short side of the mating market should enjoy more surplus from their presumably greater bargaining power (Becker, 1973). Moreover, when women are scarce, men should compete harder to increase their mate value. Positive assortative matching predicts that high-income women, in particular, benefit when the income of high-income men increases. However, we conceptually and empirically show that if women generally prefer men who possess not only a high-income, but specifically higher income than themselves, that is, a reference-dependent preference (RDP) for mate income, then high-income women can be worse off in the marriage market when there are more men or when men are richer. Such a preference was suggested by Ong and Wang (2015) with Chinese online dating field experimental data, Hitsch, Hortaçsu, and Ariely (2010b) with US online dating empirical data, and corroborated by Bertrand, Kamenica, and Pan (2015) for married couples with US household survey data. We outline a standard microeconomic basis for both the gender difference in preference for mate income and women's RDP in the literature review section.

The key insight for the comparative statics implications which we test in this study is that, with an RDP for mate income, increases in a woman's income reduces the pool of the men she most prefers, while expanding the pool of other women who most prefer these same men. In the context of an RDP for mate income, the fierceness of the competition a woman faces for the men she most prefers "escalates" as her income increases. The main focus of this present study is how this escalation in competition can, moreover, be exacerbated by increases in the income and availability of high-income men. Either may boost the expected return of pursuing such men: the former increases the value and the latter increases the probability of getting such a man. The *direct effect* of increases in the income and the availability of high-income men benefit high-income women. In the case of increases in men's income, men are more desirable. In the case of increases in the local sex ratio (number of men/number of women in a city), more high-income men are available for each woman to desire. In either case, the higher *ex-ante* expected returns for pursuing these high-income men may also increase the number of low-income women (particularly the beautiful among them) who might switch from pursuing low-income men to pursuing these high-income men as well. A greater number of women can, therefore, desire the same high-income men. Accordingly, an *indirect effect* of both increases in the income and availability of high-income men is the increased "entry" of low-income women into the matching market for high-income men. As a result, high-income women, who are averse to matching with low-income men, are worse off in the marriage market. The indirect effect is likely to dominate the direct effect for high-income women, whereas the opposite is true for low-income women, who can be satisfied with matching with low-income men. Consequently, high-income women can on balance be worse off in the marriage market relative to low-income women or even in absolute terms when high-income men are even richer or more plentiful as a result of this 'competitive entry' of low-income women into the market for high-income men. Such may be the situation in China, where both the sex ratio and men's income compared with women's have been increasing dramatically (Ge & Yang, 2014).

Facial beauty can induce further differentiation in the search behaviour of women within an income level because facial beauty may affect their prospects of attracting a high-income man. Beautiful low-income women may, therefore, be more likely to enter the market for high-income men when those men are more plentiful or richer than plain-looking low-income women. The plain-looking low-income women, rather than face the increased competition for high-income men, may be more willing to settle for a potentially better selection of low-income men. Given women's RDP, such differentiation in search behaviour by beauty among women should decrease as their income increases.

We exploit variation in the local sex ratio and the incomes of men across Chinese cities to test for our competitive entry hypothesis. We use three datasets: field experimental data with random assignment of income, the Chinese Census, and the China Family Panel Studies (CFPS) household survey data, which happens to contain beauty ratings of those surveyed. The local sex ratio of a city within a certain age range can be regarded as representing the *ex-ante* prospects for each side of finding a marriage or remarriage partner, and thus, as a measure of the competitiveness of the mating market (Becker, 1973). For the online dating field experiment, we chose 15 major cities for variation in local sex ratios and measured the variation across these different types of women's *relative* search intensities for men with different income levels. In this experiment, we randomly assigned three income levels to 450 artificial male profiles on a large online dating-for-marriage website and recorded the incomes and other characteristics of 1,811 "visits" from women to these male profiles. The women visitors were divided into high-, medium-, and low-income levels. Moreover, we also had a random sample of nearly two-thirds of these female visitors' profile photo rated for their beauty.

Consistent with our competitive entry hypothesis, we show that the search intensity of the beautiful low-income women for high-income men increases with the local sex ratio and the income of high-income men. Consistent with exiting in the face of greater competition for high-income men and substitution towards low-income men, the search intensity of the plain-looking among the low-income women for high-income men decreases with the same. Likewise, consistent with efforts at entry deterrence on the part of high-income women, the search intensity of high-income women—even the plain-looking among them—for the high-income men increases with the local sex ratio and the income of these high-income men. The lack of exiting on the part of the less beautiful among high-income women is expected if women's RDP for mate income makes them averse to settling for low-income men even when the competition for high-income men increases.

Our analysis of the Chinese Census data finds the expected ultimate consequence of the increased entry of beautiful low-income women into the market for high-income men. High-income women's probability of marriage decreases relative to that of low-income women as the local sex ratio and the income of high-income men increase, notwithstanding the uniform (with respect to their beauty) increase in these women's search intensity for these men. By contrast, low-income women's marriage probability weakly increases with the local sex ratio and on the income of high-income men, despite their aggregate (irrespective of their beauty) search intensity for these men not increasing with either. Although the negative effect of men's mean income on high-income women's probability of marriage may in part be due also to those men's marginal utility for beauty increasing with their income, we can find no standard explanation for the negative effect of the sex ratio on high-income women's probability of marriage relative to low-income women, especially in the context of the generally accepted assumption of positive assortative matching of spouses on income. Moreover, consistent with our hypothesis of competitive entry by beautiful low-income women into the market for high-income men, the CFPS data suggest that it is the plain-looking high-income women's probability of marriage which decreases not only relative to plain-looking low-income women but also in absolute terms as the local sex ratio increases. In contrast, the probability of marriage of the beautiful high-income women weakly increases relative to beautiful low-income women.

We also find evidence supporting prior findings that rising male wage inequality is a significant contributor to women delaying their marriage (Gould & Paserman, 2003).[1] Indeed, we control for the mean and standard deviation of men's income in all regressions. Given these controls, we show with Census data that only high-income women experience a decreased marriage rate when high-income men's income increases, while low-income women's marriage rate increases. We provide an additional treatment variable, the local sex ratio, which we show contributes to the divergence of marriage rates among women of different income levels with Census data and among different income and beauty levels with household survey data. We, moreover, provide novel data on the search behaviour of women and show that this search behaviour varies with sex ratio according to the income and beauty level of the women, which given our RDP framework, predicts the divergence in marriage rates we find with Census and household survey data.

Our findings would be less surprising if the deterioration in the marital prospects of high-income women can be attributed to the cross-border migration of brides from low-income regions (Weiss, Yi, & Zhang, 2017). In fact, significant internal migration has occurred within China in recent years. However, because the local sex ratio includes the migrant population, the greater influx of low-income women migrants, which *reduces* the sex ratio, cannot be the driving factor for the association between a *higher* sex ratio and the decreased rate of marriage among high-income women. Moreover, the sex ratio is controlled for when we find that high-income women are adversely affected by increases in the income of high-income men.

Instead, our online dating evidence suggests a shift in the attention and search efforts of low-income women towards high-income men within cities when these men are richer or more plentiful. In this sense, the topic of our study is also related to migration, but of attention, rather than across borders.[2] This shift in attention in matching markets is itself becoming an area of study in the nascent theoretical matching literature on directed search

---

[1] Despite the novelty of our findings for high-income women, our empirical results support standard theories – when we average across women of all income levels and beauty. Consistent with more outside-options from the greater availability of men, the marriage probability of women *on average* increases with sex ratio.

[2] Attention has been shown to be sensitive to the expected surplus in the price dispersion literature (Morgan, Ong, & Zhong, 2018).

(Chade, Eeckhout, & Smith, 2017). However, to our knowledge, no existing search model captures the rich phenomena documented here. It is beyond the scope of this study to fill in the theoretical gap. Rather, we focus on the effect of the sex ratio and men's income on the search *effort* levels of women with different income types and on these women's probability of marriage, rather than the search behaviour itself. Such effects are well captured by a simple game theoretical contest example that we provide in Appendix 1. It exhibits a contest game between high- and low-income women for high-income men prizes. The example shows that high-income women can be negatively affected in the marriage market by the increased mate value of high-income men when low-income women decide to compete for them instead of settling for their low-income men option. It also shows that low-income women's effort to win over the high-income men is more influenced by their beauty than high-income women's by theirs.

II. Related Literature

We build on Hitsch et al.'s (2010a) framework for analysing first-contact email behaviour to reveal user mate preferences. As they point out, the Gale-Shapley model, in particular, the deferred-acceptance algorithm approximates the behaviour of online daters in their exchanges of signals of interest and emails if the search frictions they face are negligible. Hitsch et al. (2010a) show that preferences revealed in the online dating context are correlated with preferences revealed in actual marriages. We focus on the click-throughs of dating profiles on preliminary search results. We argue that such click-throughs, which are prior to and necessary for making first-contact emails, are also reflective of dating preferences.

Our conceptual framework of escalating competition is similar to other theories in predicting that women's marriage rate decreases with their income (Isen & Stevenson, 2010) or educational attainment (Boulier & Rosenzweig, 1984). Recent research has focused on the potential effect of social norms in the US (Bertrand et al., 2015), Asia (excluding China) (Hwang, 2016), and other developed countries (Bertrand, Cortes, Olivetti, & Pan, 2016). We too do find women's marital prospects decrease with their educational attainment, but not with income, once educational attainment is controlled for. However, entirely different from the prior literature, we examine the comparative statics effect of the local sex ratio and men's income on the search intensity and the probability of marriage of high- and low-income women.

Our study is closely related to but still differs significantly from the burgeoning literature on the effect of the local sex ratio on the competition for mates, particularly in China. Evidence exists for the expected increase in competition among men or their supporting families and the relaxation of competition among women. The rise in the local sex ratio predicts not only increases in men's level of entrepreneurship (Yuan, Rong, & Xu, 2012), men's work hours in dangerous and risky jobs (Wei & Zhang, 2011), male criminal activities (Edlund, Li, Yi, & Zhang, 2013), the savings of families with sons (Wei & Zhang, 2011), the time men spend on housework within households (Du, Wang, & Zhang, 2015), and women's participation in decision making (Edlund et al., 2013). The rise in the local sex ratio decreases women's educational attainment and employment (Edlund et al., 2013). These studies on outcomes, which focus primarily on the behaviour of men, confirm Becker's (1973) theory that the bargaining position of women improves as the sex ratio rises.

The effect of the sex ratio on competition for mates has been studied in countries other than China. For example, consistent with Becker's theory, Abramizky et al. (2011) find that the low sex ratio (that is, shortage of men after World War I) led more men to "marry up". That is, the short side of the market benefited from its shortage. Our contribution to this sex ratio literature is to test for heterogeneity in the effect of the sex ratio on subgroups of the mating market. In contrast to previous studies, we demonstrate that a subgroup of the short side of the market (women in our case) was negatively affected by RDP despite their shortage, which is the opposite result and contrary to what one might expect based on Becker's theory.

Women's RDP can be explained as a consequence of the classical assumption that women generally specialize in household production after marriage due to traditional gender roles (Becker, 1973).[3] This assumption is fleshed out in, for example, Caucutt et al. (2002), which posits that women take up the bulk of the burden of childcare and will marry only if the utility from pooled consumption and investment in children in the context of marriage is higher than their outside-option of being single (consume and/or raise children using their income). Women's own income is their reference point within our conceptual framework.

With regard to women's labour market opportunity cost from specialization in household production, recent evidence supports the implication that Becker draws from the assumption that women specialize in household production after marriage, which is that marriage rates would increase with the gender gap in wages because the gap decreases women's labour market opportunity cost from specialization in household production (Autor, Dorn, &

3 Women's RDP has also been documented in the large sociological literature on female hypergamy.

Hanson, 2018; Shenhav, 2018). Recent evidence also confirms Becker's prediction that marriage and childbirth decrease women's labour market participation. Women in the West (Lundberg & Rose, 2000) and in China (Feng, Hu, & Moffitt, 2017; Hare, 2016) often relinquish full-time work after marriage and childbirth. Among these women, the more highly educated are relatively more likely to "opt-out" completely (Hersch, 2013). Women's decreased labour market participation after marriage, particularly those in the high earning professions in the US (Goldin, 2014), and childbirth in Denmark (Kleven, Landais, & Søgaard, 2018; Lundborg, Plug, & Rasmussen, 2017) and Norway (Bütikofer, Jensen, & Salvanes, 2018), is the main source of gender differences in wages. Furthermore, women anticipate and report preferring (Parker & Wang, 2013) decreased labour market participation after marriage and childbirth in the US, even before they marry or graduate from college. Women also anticipate a relatively higher income husband (Wiswall & Zafar, 2016), who can potentially support their premarital standard of living. Correspondingly, men anticipate a lower income wife and no change in labour market participation after marriage.

Women's anticipation of lost income after marriage helps explain the asymmetry in preference for mate income found in Ong and Wang (2015). To the extent that a woman's income is expected to be forgone after marriage, her income will not benefit her potential husband. Men's apparent lack of income-based attraction to women as marriage partners may reflect their anticipation of the women's loss of income. Given women's likely and anticipated reduction in income after marriage, it is natural for them to seek out a husband whose income would substantially offset the opportunity cost of their *potentially* decreased labour market participation and specialization in household production to maintain their standard of living.[4] We provide a conceptual framework to understand and empirical evidence for the intra-gender competition between women resulting from their search for such husbands. Moreover, we provide evidence consistent with the dominance of women's RDP for a higher income mate over a possible men's RDP for a lower income mate.

Our empirical results appear to be novel in the context of standard search and matching theories. The sex ratio within Becker's theory of marriage matching determines who gets

---

[4] Such behaviour is in-line with job search behaviour in general which has been found to be reference-dependent on the prior job (Dellavigna, Lindner, Reizer, & Schmieder, 2017). In many cases, women literally also do search for jobs with less demanding more flexible hours, travel, and consequently lower pay after marriage and childbirth (Blau & Kahn, 2017). Men, in contrast, do not. This notion that a woman may search for a mate with a view to offsetting her opportunity cost is consistent with the long standing theory of habit formation and with the recent behavioural theory of reference-dependent preference in which the reference point is lagged consumption (Kőszegi & Rabin, 2012). Regardless of whether RDP has a standard microeconomic basis, we consider RDP as a primitive notion and focus on the potential over-entry/congestion effect of a targeted search for a mate based on income. In particular, we test for the possibility of crowding out of high-income women by the competitive entry of low-income women, when high-income men become more plentiful or richer.

married. In particular, if the sex ratio is higher than one, some men are not married. Among married couples, the sex ratio also determines who gets the marital surplus in a transferable utility (TU) framework. Within a TU framework, when the sex ratio is higher than one, wives get the entire marital surplus.

Although we provide evidence that high-income women are negatively affected in the marriage market by increases in the sex ratio or high-income men's income because they exert greater search effort and have a lower probability of marriage, the welfare effects remain ambiguous. High-income women's "plight" results from them being rationally more selective about whom to marry because their singlehood outside-options are better than that of low-income women. Our focus in this study is not to compare observable traits and infer marital surplus as is done in the standard marriage literature. Accordingly, we leave for future research the development of theories to demonstrate that women's RDP may support the high correlation of other observable traits such as education in couples. Our focus is rather the overlooked comparative statics effect of intra-female competition for high-income men. Moreover, although we also provide evidence that such competition is motivated by the women wanting to replace their lost income after marriage and childbirth, we do not directly observe this motivation.

III. Online Dating Field Experiment

Experimental Design

We outline the design of our experiment as follows. We use the experiment to identify the comparative statics of women's preferences for mate income as the sex ratio and men's mean income vary. The analysis of this experimental data forms the basis of our predictions for marital patterns observed in the Census and household survey data in subsequent sections.

Our field experiment extends Ong and Wang (2015) by testing for women's preference for mate income across many cities that vary in their local sex ratio. Our experiment is in the tradition of the considerable literature on correspondence studies of labour market discrimination. We used one of the largest online dating websites in China, with a reported membership of 100 million members in 2016. The website we used (and its competitors) advertises itself as a marriage matching website for white-collar professionals between the ages of 25 and 45 years of age.

The users of this website can create a profile for free. The profile must include demographic (e.g., age and gender), socioeconomic (e.g., income), and physical characteristic (e.g., height) information, at least one photo, and a free-text personal statement. These requirements are standard to most online dating websites. Users may also add more information, and in particular, verifiable information to increase the “credibility” of their profile.[5] Users can browse, search, and interact with other members after registration. Generally, users start by entering their preferred age range and geographic location of partners into the search engine. The query returns a set of abbreviated profiles which include: ID, picture, nicknames, age, city, marital status, height, the first two lines of a free-text statement, and perhaps unique to China: income. Users can then click a link and “visit” the full profile, where they can signal interest for free. Emails, however, require membership. The membership fee was 10 CNY/month at the time of the experiment when 1 USD was approximately 6 CNY. We recorded only visits.

We constructed our 450 male profiles on this website by collecting nicknames, pictures, and statements from real profiles from *another* website that would have automatically hidden them after a month of inactivity.[6] These profiles were posted for only 24 hours, after which, the accounts were closed. To further minimize any possibility of being recognized by acquaintances, we ensured that their picture was assigned to a province (city) that was different from their work area or birthplace.

We assigned 30 profiles of five ages: 25, 28, 31, 34, and 37; three incomes: 3-5, 8-10, and 10-20 (1*k* CNY) per month, which we will call low- (*L*-), middle- (*M*-), and high-(*H*-)income, respectively; and two replicas to each of the 15 major cities (see Appendix 2), resulting in 450 profile “slots.” Then, we randomly assigned 450 pictures, nicknames, and personal statements to these 450 slots. For the profile’s fixed traits, we gave all male profiles the height of 175cm. Birthdays were within eight days of each other and of the same zodiac sign. All of our profiles listed college education and the marital status of “single with no children”

[5] The credibility of the profile is indicated by a positive score, which can be increased with additional forms of verification, e.g., government-issued identification. All of our profiles simply display phone verification and one photo, giving them the minimal score. Such scores would not generally affect visit rates because they do not appear in search results. To affect visits, users must search specifically for low-credibility profiles. Even then, such searches would not affect visit rates across our profiles. The across-profile visit rates are the basis for our findings.

[6] We are unaware of legal restrictions on the non-commercial use of user created content uploaded to social media websites in China. We assumed that such restrictions, if they exist, are weaker in China than in the United States, where our research activities would also fall under the “fair use” exemption to the US copyright law. Major US social media websites explicitly announce terms of use that effectively make uploaded user created content public domain. For example, see, “publish content or information using the Public setting” in https://www.facebook.com/legal/terms.

Visits to our profiles are likely to be brief, as they contain no information beyond what was already revealed in the search engine results. In fact, no one pursued further contact with any of our profiles. Our profiles are spread out among many other profiles on any given day. They are also spread out across many days. Users of this website are unlikely to encounter our profiles more than once (if at all).

Chinese universities, similar to their European counterparts, do not have IRBs to approve the ethics of experiments. However, to the best of our understanding, our design falls under the “minimal risk” exemption from IRB approval. “Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.”

See for example: http://humansubjects.stanford.edu/hrpp/Chapter9.html and "buy a house after marriage" (that is, did not own a house).

Users can see our profiles' picture, nickname, age, city, marital status, height, income and the first few lines of a free-text statement in their default search results. They can then click a link and visit the full profile, which contained no additional information. For each of our profiles, we can see the profiles of the visitors by clicking their link in the history of visitors. The website records only visits to individual profiles once from any visitor even if they make multiple visits. Although visits across different profiles need not be from unique visitors, in our data, all visits were from unique visitors. This lack of repeated visits is expected because we took care to distribute our profiles between many other user profiles. Thus, to our knowledge, each of our data points is from a unique visitor. In any case, random assignment of characteristics to our profiles should rule out the individual idiosyncratic factors of our visitors as the primary driver of our findings.[7]

We created profiles the day before to allow the website time to register them. Profiles of each age, income, and city combination were equally distributed in 12 days, with 35 to 40 profiles each day. We randomly logged in these 35 to 40 profiles with at least five minutes between any two, extending to at least 10 minutes between any two profiles in the same city. This procedure left at least one page between each of our profiles. For the 12 experimental days from August 23 to September 3, 2014, each account was open for only 24 hours. We alternated between logging in the next day's profiles and collections of data on the previous day's visit data. The total login/collection time was three to four hours per day depending on computer speed and the total number of visits our profiles received.

In total, our male profiles received 1,811 visits from women. 1,474 of these visits have photos. Among these, 1316 were of a quality useable for rating, e.g., of high enough resolution, had faces not obscured by sunglasses...etc. We had a random sample of two-thirds or 867 of these women visitor's photos rated for their beauty using a proprietary rating program accessible through a standard web browser. In the rating program, each female visitor's photo (i) is randomly matched with 10 other photos ($j \neq i$) from the pool of all photos. Each photo is selected with replacement from the pool of photos 20 times. Each photo was on average rated 200 times, which is approximately 10 times the frequency of other studies. A total of 692 Chinese raters (326 male) rated these 867 photos. The raters were graduate students from the Peking University HSBC Business School recruited through

---

[7] The pictures, nicknames, and the first two lines of the personal statements were randomly assigned to profile slots. If the women's choices were based on anything other than the income of the male profiles, we would find a uniform distribution of clicks across incomes and cities.

a mass email. We used two rounds for rating (one-third of photos in each round), because of our limited capacity to recruit raters during the first-round.[8]

We asked raters to choose the more physically attractive within each pair of 100 pairs instead of asking for a numerical rating within a specific range of numbers, as is standard in the field (Hamermesh & Biddle, 1994). This binary judgment may be easier and more precise than assigning a number to how good-looking someone is based on a numerical scale. The binary decision also avoids potential scale differences across individuals and genders which would add noise to our data. The software then aggregates the ratings for each photo into a continuous number between 0 percent, for the least attractive, and 100 percent, for the most attractive. For each photo, these numbers represent the share of other photos that the raters on average found less attractive.

We also use data from another experiment which was run simultaneously with 390 female profiles in the same 15 cities. These female profiles had ages of 22, 25, 28, 31 and 34, a height of 163 cm, were college educated, and had incomes of 5*k* to 8*k* CNY/month. For our main experiment, we utilize the reported incomes of the male visitors attracted by these female profiles (restricted to the same ages (25-37) as our male profiles) to construct the distribution of men's income on the website in the 15 cities.

Description of the Data

The summary statistics of age, income, and education for each gender of our visitors are in A-Table 6 and A-Table 7 of Appendix 2. The women visitors range in age from 18-45. In 2005, the median age of first marriage for women was 23 according to the 2005 1 Percent Population Survey (often called "the mini-Census"). 98.5 percent were married by the age of 30. The 2010 Census does not contain micro-level data of individual characteristics such as income or education. However, the CFPS data for 2010 show a similar pattern to the 2005 mini-Census. Accordingly, we find an abrupt decrease in the share of searches of women of all income levels, especially high-income women at age 30 in A-Figure 2.[9] We focus throughout the paper on women in the age range of 20 and 30 years old. We use the visits from women age 31 and 45 as a check whether the sex ratio we use does not affect the women outside of the age group that is the focus of our study.

---

[8] We paid raters 5 RMB to rate 100 pairs of photos in the first-round (January 4, 2016) and in the second-round (November 23, 2016). Given the few minutes it took to rate all 100 photos, our payment was relatively high for China. We set a high wage to attract sufficient numbers of raters within a short period.

[9] We find no such decline at any age for men.

We calculate the local sex ratio (the *log* of the number of men/number of women) using county-level data based on the full sample of the 2010 Census.[10] The local sex ratio for this experiment is defined as the number of males between the ages of 22 and 32 years old over females between the ages of 20 and 30 years old at the time of experiment in 2014, proxies by males between the ages of 18 and 28 years old and females between the ages of 16 and 26 years old in the 2010 Census, which allows for the standard two-year age gap observed between married couples in China.

The individual characteristics of our women visitors that we collected include income, age, years of education, and height. The website allows for the reporting of only nine income levels (<1, 1-2, 2-3, 3-5, 5-8, 8-10, 10-20, 20-50, and >50 in 1*k* CNY). We define *h*-, *m*-, and *l*-women by absolute cut-offs: *l*-women, that is, <3*k*/month, *m*-women: 3-5*k*, *h*-women: >5*k*. We set these levels to approximate the bottom-, middle-and top-thirds of women, which are the grouping we use throughout the paper. For women between the ages of 20 and 30 years old, the shares of *l*-, *m*-, and *h*-women are approximately 25, 39 and 26 percent of our visits, respectively. The low-income level is the omitted benchmark in all regressions.[11]

In the following analysis, we show that the treatment effects of the sex ratio and men's income in our online dating sample are consistent with the treatment effects in the representative samples in the Census and CFPS datasets. For the comparison to be valid, we need to rule out that the consistency in treatment effects is due to a fluke of comparing disparate populations that happen to have similar treatment effects. We cannot directly compare the women visitors' characteristics to the women in the Census data because the 2010 Census does not contain individual-level income data.[12] (The 2015 mini- census is not currently available.) We also cannot make the comparison using the CFPS data because that dataset covers provinces and has minuscule samples for the cities used in our experiment. Instead, we compare the characteristics of our female visitors to the single women in the Urban Household Survey (UHS). The UHS is conducted by the National Bureau of Statistics and covers a representative sample of the urban population at the city-level in China. The UHS 2012 sample is the closest year to 2014 (the year of our experiment)

[10] See the tabulation of the 2010 Population Census at the county level by the National Bureau of Statistics. The 2010 Census released only the aggregate number of people of each gender in five-year age groups: 20-24, 25-29, 30-34, 35-39, and 40-44. In our calculation of the local sex ratios, we assume each age within the five-year age group has same population size, e.g. the population size of age 22 is 1/5 of the population size of age group 20-24, which is reported by the Census.

[11] We used absolute cut-offs for incomes in the online dating section of the study because the website aggregates incomes into nine levels.

[12] We also cannot make the comparison with the 2005 mini-census, which does contain individual income data, but is too distant in time to be comparable, especially given the fast pace of economic growth in China.

available. It covers four of the cities used in our experiment, namely, Dalian, Shanghai, Guangzhou, and Chengdu. These cities are located in the Northern, Eastern, Southern, and Western China, respectively.

In comparison to single women in the UHS dataset, our female visitors are approximately four years older. 55 percent of single women in the UHS sample are younger than 26, whereas 26 percent of our female visitors are within that age range, with half concentrated in the age range of 26 and 32. We find, similar to Hitsch, Hortaçsu, and Ariely (2010a) in the US, that the website users in China are more educated (one year, in our case). In comparison to the UHS sample, 14 percent more of our women visitors have a college or higher level of educational attainment. (These distributional results are available on request.) Not surprisingly, given their older age and higher educational attainment, our women visitors also have a higher income than the general population.

Hitsch et al. (2010a) find that differences in educational attainment and income become insignificant after controlling for internet use in the US. Controlling for internet use, which has been growing rapidly in China, would likely have the same effect. However, we cannot make the same adjustment because the UHS does not contain internet usage data. In the case of China, one additional potential reason why the website users' educational attainment and income are higher is that educational attainment and income increased rapidly between 2012 and 2014 as a result of the Chinese Government's expansion of educational opportunities (Knight, Deng, & Li, 2017) and China's characteristic rapid economic growth. Instead of controlling for internet usage, we perform a Mincer-type regression of income on age and educational attainment for our women visitor sample and the single women in the UHS sample and plot the distribution of the residuals of each regression in Figure 1.

[Insert Figure 1 here]

These two distributions almost entirely overlap. The remaining differences in the distribution may be attributable to the limited sample size of 435 visits for our data for the four cities compared. We, moreover, directly compare the effect of age and educational attainment on income for the two samples in a dummy variable regression using the combined UHS and experimental samples in Table 1.

[Insert Table 1 here]

None of the coefficients for the interactions between the dummy for the experimental sample and the age and educational attainment variables are significant from the single women in the UHS sample. The overlapping of the distributions in Figure 1 and the lack of significance of coefficients for the interactions in Table 1 suggest that our women visitors are not a specially selected sample of the general population, especially after controlling for age and educational attainment, which we do in all subsequent regressions.

Graphical Analysis

Before we present our main findings, we first establish graphically in Figure 2 the correlation between the sex ratio and men's income on the website and in the surrounding city for the cities used in the experiment.

[Insert Figure 2 here]

The upper panel shows the distribution of men's income on the website and in the surrounding city for the 15 cities used in the online dating experiment. The upper leftmost panel shows the distribution of reported incomes for our entire sample of 5,535 visits from men between the ages of 18 to 45 years old from the other experiment. The upper middle panel is restricted to the 3,520 visits from men between the ages of 25 to 37 years old, which is the age range that matches our male profiles. The upper right panel is restricted to the 2,832 visits from men between the ages of 22 to 32 years old, which is the age range of the men that we focus on throughout this paper. The lower left panel shows the income distribution of the men between the ages of 22 to 32 years old within the same 15 cities. The lower right panel shows that for the larger sample of 57 cities for which the 2005 mini-Census contained more than 300 respondents of either gender in the age range of 20 and 30 for women and 22 and 32 for men.[13] These 57 cities will be the focus of our analysis of women's marriage probabilities.

Within each panel, the cities are divided into top-, middle-, and bottom-five city groups in terms of the magnitude of the local sex ratio. For the online dating data in the two upper panels, the sex ratios are defined as the number of males between the ages of 22 and 32 years old over the number of females between the ages of 20 and 30 years old at the time of the experiments in 2014, based on the 2010 Census.[14] The sex ratio in the lower two panels is similarly defined (with a minimum sex ratio of 0.79 and a maximum of 1.13) but use data from the 2005 mini-Census.

The upper panels show that our male visitors' income distribution of cities with high sex ratio (top-five-city group) is noticeably more right-skewed than that of those in the medium-, and bottom-five-city group for every age group. For the 2005 mini-Census sample in the

[13] The 2005 mini-Census is the last available with income data.

[14] Data from the 2015 mini-Census are unavailable. The Chinese Government has scarcely made any micro data available in the last five to 10 years.

lower panels, we observe a slightly greater right skew for the distribution of men's income for the middle-third sex ratio cities rather than the top-third. These panels suggest that male online daters in the 15 cities of our experiment and the men in the same cities and for the 57 cities for which we have sufficient sample size are not poorer in higher sex ratio cities. The regression results for the 57 cities in A-Table 4 and A-Table 5 in Appendix 2 further suggest that the correlation between the local sex ratio and men's income is not negative. These results imply that increases in the sex ratio are not driven by a disproportionate increase in the share of low-income men compared with high-income men. Therefore,

Observation 1. Men in higher sex ratio cities are richer on average than in low sex ratio cities. Higher sex ratio cities do not have a greater share of poor men either in the city itself or on the website associated with the city.

The increase in women's visits to high-income male profiles as sex ratios increase, which we hypothesize as possible, is already evident in the graphs of our data in Figure 3.

[Insert Figure 3 here]

The graphs in Figure 3 exhibit visits by women to male profiles in three groups of cities, that is, top-, middle- and bottom-five, based on the local sex ratio, defined as the number of males between the ages of 22 and 32 years old over females between the ages of 20 and 30 years old at the time of experiment in 2014. These populations are proxied by males between the ages of 18 and 28 years old and females between the ages of 16 and 26 years old in the 2010 Census, which allows for the two-year age gap observed between married couples in China. The horizontal axis indicates the ages of our male profiles.[15] The vertical axis displays the percentage of visits (%), which is the total number of visits received by each type (age and income) of male profiles in each five-city group divided by the visits to all our profiles over all male income types in the same five-city group. Although the gap between the graphs of the visits of women to low- and middle-income men does not show a clear pattern, the gap in the graphs of women's visits to middle-to high-income men and from low- to high-income men clearly do increase on local sex ratios. The pattern suggests that the marginal impact of increases in men's income on the visits of women increases with the local sex ratio.

Recall that we fixed the number of profiles (10 for each of our three income levels) across all cities in our online dating experiment. Thus, our high-income profiles should have received a constant share (relative to our medium- and low-income profiles) of all visits in the higher sex ratio cities given a constant distribution of visits to the three income levels across all cities, not a larger share, as our main findings indicate. In other words,

Observation 2. Our high-income profiles can only receive a larger share of all of the visits to our profiles in higher sex ratio cities if women visit high-income men more than low-income men when men are richer or plentiful.

To see which group of women contributed most to the increased visits, we grouped female visitors into three income levels: <3, 3-5, and >5 (in 1*k* CNY) in Figure 4.

[Insert Figure 4 here]

These levels are labelled as *l*-, *m*-, and *h-women*, respectively and are represented by three lines. These three income levels for women are lower than the three incomes levels for our male profiles, because as in most countries, women in China typically earn less than men. The left (low sex ratio cities) and the right (high sex ratio cities) panels of Figure 4 show that women of all income levels visit high-income male profiles with greater probability. However, our focus here is not on the mate attraction effect of the absolute level of men's income on women's behaviour, but rather on the effect of men's income relative to women's incomes on women's behaviour in three respects.

First, each panel shows that the slopes of the lines connecting the mass points of these probability mass functions rotate counter-clockwise. This rotation indicates that the probability of women's visits to high-income male profiles increases with their reported incomes. Second, we show a kink in the graph of the high-income women at point B and B', which suggests that their visit rates to high-income male profiles (10-20*k*) increases significantly compared with that of middle-income male profiles (8-10*k*), that is, as the profile's income exceeds the women's average income (which is approximately 15*k*). Third, previewing the main findings in Table 2, Figure 4 displays a further counter-clockwise rotation from the left to the right panels (AB – BC and A'B' – B'C') from the cities with low sex ratio to those with high sex ratio for high-income women. For example, although a small percentage of *h*-women's visits were to male profiles with reported earnings of 3-5*k*/month in the bottom-eight sex ratio cities (point A), visits to such male profiles remain visibly lower to the point of being nearly zero in the top-seven cities (A'). Approximately 75 percent of *h*-women visits were to the 10-20*k* male profiles in bottom-eight sex ratio cities (C), but approximately 85 percent of their visits was to that type of profile in the top-seven cities (C'). These three levels of evidence suggest the increased search effort of high-income women, due to women's RDP, even before we impose controls econometrically.

Regression Analysis

Here, we discuss the formal analysis of search behaviour as revealed by our online dating data, which constitute the basis for our predictions for married couples observed in the Census and household survey data. We exclude 51 visits without income information from the 1,811 visits we received from women, leaving 1,760 visits for analysis. Each of our 450 male profiles is at one of the three income levels in one of the 15 cities of the experiment. Let the income level of the male profiles that woman i chooses to visit be represented by the latent variable y_i^*. We observed her visits if these were made to one of our three income types of male profiles. We treat each as one of the three choices in an ordered probit model

$$y_i^* = X\beta + \varepsilon_i \qquad \text{Eq.(1)}$$

where X includes *m-women dummy* (medium-income women), *h-women dummy* (high-income women), and *log* sex ratio (the *log* of the number of men/number of women—sex ratio from this point forward) and its interactions with the above two dummies, and individual and city characteristics.[16]

We control for the wage distribution (means and standard deviations of men's and women's incomes) from the incomes of our visitors, and therefore, women's average opportunity costs for presumably specializing in household production after marriage by testing for the change across cities in their search intensities. We use the ordered probit regression to model the probability that a woman from a specific income level visits a male profile of a specific income level among all income levels of male profiles. We interpret this probability as the search intensity for a man of a specific income level, which being a probability, is normalized by the total number of visits per women's income level at the city-level.

[Insert Table 2 here]

Table 2 displays the results of the ordered probit regression of women's visits as a function of their income and the local sex ratio. The positive term for the *h-women dummy* (0.502) in column (1) indicates that high-income women visit high-income male profiles more than low-income women do, thereby confirming our impression from Figure 4 and supporting previous findings (Ong & Wang, 2015). Column (1) of Table 2 demonstrates not only that the intercept for high-income women is higher than that of the benchmark low-income women,

---

[16] Note that our treatment variable in the experiment is men's income type (*H, M,* and *L*). However, this may not be evident in our ordered logit regression because men's income type does not appear on the RHS. Nevertheless, this information is implicit in our dependent variable, that is, the *log* odds of visiting higher income men.

but also that the difference increases with the sex ratio (1.765). In apparent contradiction to our competitive entry hypothesis, the coefficient for *sex ratio* for the benchmark low-income women is small and statistically insignificant (0.432) in column (1).[17] However, the lack of significance can also be attributed to low-income women, who enjoy more outside-options among low- and medium-income men. We show that the lack of significance is due to the less attractive among low-income women decreasing their search intensity for high-income men. This decrease exerts an offsetting effect on the increased search intensity of the beautiful low-income women for these same men.

Facial beauty is generally regarded as an important characteristic for females because facial femininity, which adds to female facial beauty, signals high levels of the female hormone oestrogen, and therefore, fertility (Rhodes, 2006). However, with few exceptions, facial beauty is neglected in the literature on the economics of marriage. We focus on the effect of facial beauty in the mating market in a companion study. In this article, we used data on facial beauty to allow further differentiation in the search behaviour of women within an income level because facial beauty may influence their prospects of attracting a high-income man. We argue that this difference in prospects can be modelled as a difference in the cost of competing in a standard contest model in Appendix 1. Therefore, we control for the beauty percentile ranking of female visitors in column (2) of Table 2 which we acquired for a random sample of two-thirds of these visits.

[Insert Figure 5 here]

Figure 5 shows that the mean and standard deviations of women's facial beauty do not vary systematically with the sex ratio of the cities used in the experiment for any of the high-, medium- and low-income levels of women. Importantly for our previous finding of an insignificant increase in the search intensity of low-income women for higher income men, column (2) of Table 2 also reveals heterogeneity in the reactions of women according to their beauty rank and income level to increases in the sex ratio. The significantly negative coefficient for *sex ratio* (-6.219) indicates a pronounced decrease in the search intensity among of the benchmark plain-looking low-income women for high-income men when the sex ratio increases. By contrast, the highly significant positive coefficient for *sex ratio*beauty* (12.131) indicates a pronounced increase in the search intensity among the

[17] This insignificance can be due to our design being naturally biased towards a negative effect for increases in sex ratio. Our fixed number of high-income profiles at fixed income levels should receive fewer visits in cities with higher sex ratio, where men on the website (and in the surrounding city) are richer and more plentiful. Hence, the reader should perhaps interpret the weakly negative coefficients as weakly positive.

beautiful low-income women for high-income men when the sex ratio increases. Thus, increases in the sex ratio induce divergent reactions among beautiful and plain-looking low-income women, which helps to explain the apparent lack of reaction of low-income women in aggregate (when we do not disaggregate by beauty) in column (1).

The medium- and high-income women show diminishing contrasting reactions by their beauty rank as the sex ratio increases, due to their RDP. The significant positive coefficient (5.401) for the interaction between the sex ratio and the medium-income women dummy suggests that the plain-looking among them react less negatively than low-income women to the increase in the sex ratio. In contrast, the weakly negative coefficient (-8.339) for the interaction among sex ratio, beauty rank, and the medium-income women dummy suggests that the reaction of the beautiful among medium-income women to the sex ratio is less influenced by their beauty than those among low-income women. The positive and significant coefficient (10.879) for the interaction of the sex ratio and the high-income women dummy suggests that the plain-looking among the high-income women search more intensively for high-income men when these men are more plentiful. Similarly, in contrast, the negative coefficient (-19.219) for the high-income women suggests that their reaction to the sex ratio is less positively influenced by their beauty compared with that for low-income women. This pattern of decreasing differentiation in search intensity for high-income men by their beauty rank, as the women's income level increases, is expected because the lower the women's income level, the larger their set of options among low-income men, and therefore, the greater their latitude to avoid the increasing competition for high-income men. Moreover, as indicated in column (3), these contrasting behaviours between beautiful and plain-looking women with different income levels become more significant for women between the ages of 20 and 30 years old and insignificant or reversed (-20.534) for women 31-45 in column (5), which suggests that these results are driven by women in the prime marriage market age in China. We largely confirm these contrasting behaviours for women between the ages of 20 and 30 years old in column (4) with the Bartik (1991) style instrumental variable for the sex ratio,[18] though the levels of significance of some coefficients are reduced. See A-Table 8 for

[18] We construct the Bartik-type IV sex ratio by using the well-established gender segmentation by industry (e.g., more women are employed in service industry, whereas more men in construction and manufacturing). If an industry is male- (female-) dominated, and a city has a large share of population working in this industry, then the overall sex ratio in this city tends to be biased towards men (women). By focusing on the historical industry composition of the city, we can isolate the sex ratio that is driven only by the labour demand, which is presumably not correlated with the preference in the marriage market. By using the industry-specific sex ratio at the national level excluding this city, we remove the mating preference possibly contained in the local sex ratio. The Bartik sex ratio of city c in year 2010is:

$$\bar{R}_{c,2010} = \sum_{j} \gamma_{cj,2000} R_{-cj,2010}$$

the first-stage results.

We calculate the marginal effects of the sex ratio on high-income women's probability of visits based on the coefficients of the ordered probit regression in column (3) of Table 2, where we include more controls than in column (2), keeping all variables at their mean values. A 10 percent increase in the sex ratio decreases the probability of plain-looking low-income women's (ranked in the 25th percentile in beauty rank) visits high-income male profiles by 24.2 percentage points.[19] In contrast, the same increase in the sex ratio increases the plain-looking high-income women's visits to high-income male profiles by 8.9 percentage points, which is approximately half the increase of the 16.7 percentage points of beautiful low-income women (75th percentile in beauty rank) to high-income men's profiles.

To summarize, the reaction to the increase in the sex ratio of the plain-looking among medium- and high-income women is less negative than that of the plain-looking among low-income women. The reaction of the beautiful among the medium- and high-income women is less positive than that of the beautiful low-income women. Thus, the greatest contrast between the behaviours of the high- and low-income women when the sex ratio increases is between the plain-looking high-income women and the plain-looking low-income women. The women's own beauty rank makes less of an impact on their search intensity for high-income men as the women's income increases because their reluctance to substitute toward low-income men increases with the women's own income. This result is expected if high-income women are more determined (less willing to avail themselves of the option of low-income men) to match with a high-income man, when the competition for high-income men increases, due to women's RDP.

Observation 3. Among low-income women, the more beautiful they are, the more they visit high-income men when the sex ratio increases. By contrast, plainer looking low-income women visit high-income men less when the sex ratio increases. Both effects decrease with women's income level.

After demonstrating that only the search efforts of high-income women increase uniformly,

---

$\gamma_{cj,2000}$ is the historical share of employment in industry j of city c from earlier census in 2000. $R_{-cj,2010}$ is the sex ratio of all workers employed in industry j nationwide excluding city c, in year 2010.

[19] In our ordered probit model, the probability of each type of male profile being visited is given by $P(L=1)=\Phi(\kappa_1 - X\beta)$, $P(M=1)=\Phi(\kappa_2 - X\beta)-\Phi(\kappa_1 - X\beta)$, and $P(H=1)=1-\Phi(\kappa_2 - X\beta)$, where κ_1 and κ_2 are the estimated cutoffs, and Φ is the cumulative density function of the standard normal distribution. We calculate the marginal effect on each probability's change as $\frac{\partial P}{\partial X}$, keeping all explanatory variables at their mean values. For a positive coefficient β_i of X_i, the marginal effect $\frac{\partial P(L=1)}{\partial X_i} = -\beta_i \phi(\kappa_1 - X\beta) < 0$, where ϕ is the probability density function of the standard normal distribution, and $\frac{\partial P(H=1)}{\partial X_i} = \beta_i \phi(\kappa_2 - X\beta) > 0$, wheres $\frac{\partial P(M=1)}{\partial X_i} = \beta_i \phi(\kappa_1 - X\beta) - \beta_i \phi(\kappa_2 - X\beta)$ is in general ambiguous.

irrespective of their beauty rank, when men became more plentiful, we examine the effect of the changes in the incomes of the top-, middle- and bottom-third income men (*H*-, *M*-, and *L*-men, respectively) on the website site who visited our female profiles in each city on the probability of these women between the ages of 20 and 30 years old visiting our high-income male profiles in column (6) of Table 2.[20] We omit discussion of the *M*-men because, being intermediate between *H*-men and *L*-men, the effect of the change in their mean income on high- and low-income women will be ambiguous. We also omit the results for *L*-men because they are almost always insignificant for *l*- and *m*-women, and in any case, less than 4 percent of *h*-women visit such men. We make these results available on request.

First, we control for the effect of the sex ratio. As in column (2), the sex ratio remains negative but is now insignificant. Similar to increases in the sex ratio in columns (1)-(4), the influence of beauty rank on women's response to the increase in the mean income of high-income men diminishes as women's income level rises in column (6), because of women's RDP. As with increases in the sex ratio in columns (1)-(4), increases in the income of high-income men induce opposing reactions among the beautiful and the plain-looking low-income women in column (6). When the income of high-income men increases, the significant negative coefficient of *mean income of H-men* (-0.483) indicates that plain-looking low-income women are less likely to pursue (more likely to exit the market for) high-income men. The insignificant coefficient for *mean income of H-men*beauty* (0.567) suggests that beautiful low-income women are weakly more likely to enter the market for high-income men, and strongly (0.868) more likely to enter when we instrument the sex ratio in column (7). The significant positive coefficient for *mean income of H-men*m-women dummy* (0.615) in column (6) indicates that the plain-looking medium-income women are less likely than low-income women to exit (at least weakly more likely to enter) the market for high-income men. In contrast, the significant negative coefficient for *mean income of H-men*beauty*m-women dummy* (-0.853) indicates that beautiful medium-income women are significantly less likely to enter than beautiful low-income women when the income of high-income men increases. The significant positive coefficient for *mean income of H-men*h-women dummy* (0.990) indicates that plain-looking high-income women are less likely to exit (at least weakly more likely to enter) the market for high-income men than low-income women. Similarly, in contrast, the significant negative coefficient for *mean*

[20] Recall that we gathered this income information from the men visiting our female profiles in another experiment that we conducted simultaneously with this experiment.

*income of H-men*beauty*h-women dummy* (-1.370) indicates that beautiful high-income women are less likely to enter than beautiful low-income women. These patterns are even more significant when we instrument the sex ratio in column (7). Again, this pattern of decreasing differentiation by beauty rank among women as women's income increases is expected if high-income women are more determined to match with a high-income man (and less willing to avail themselves of the option of low-income men), due to women's RDP.

A 10 percent increase in the mean income of high-income men decreases the probability of the plain-looking low-income women's visits to high-income male profiles by 13.3 percentage points. By contrast, that same increase in mean income increases the plain-looking high-income women's visits to high-income male profiles by 12 percentage points. Thus far, for plain-looking women, a parallel change in behaviour when either the sex ratio or high-income men's mean income increases. However, the parallel holds only weakly for beautiful low-income women. Beautiful low-income women's search intensity for high-income men increases insignificantly with the increase in men's mean income.

The increase in the visit rate of plain-looking high-income women to our high-income male profiles when the mean income of other profiles rises is remarkable given that our high-income male profiles are then relatively less attractive. One would expect that the change in these women's rate of visits would be negative. Given this expectation of a negative coefficient, one should perhaps interpret the lack of significance in the beautiful low-income women's visit rates to high-income male profiles, as likely significantly positive. This weaker increase in the search intensity of beautiful low-income women for our high-income male profiles is also consistent with the possibility of these high-income men exerting greater effort in their search for a beautiful girlfriend/wife when their income increases, and thereby, obviating the need of these women to increase their own search intensity for high-income men. Table 3 presents evidence that is consistent with this possibility; men's search intensity for beautiful women increases with the men's income.

These results in column (6) of Table 2 are summarized in Observation 4.

Observation 4. Even controlling for the effects of the sex ratio, among low-income women, the more beautiful they are, the more they visit high-income men when the latter's mean income increases. By contrast, plainer looking low-income women visit high-income men less when the latter's mean income increases. Both effects decrease with women's income level.

Thus far our results indicate that low-income women's entry into the market for high-income men when the sex ratio or high-income men's income increases with their (the women's) beauty rank. These results predict that the probability of high-income women (particularly the plain-looking) marrying decreases when either the sex ratio or high-income men's income increases. We test these predictions after testing for demand-side effects of men's search behaviour. We divide men into rich and poor and show the interaction between these men's probability of visits to more beautiful women and the sex ratio in Table 3.

[Insert Table 3 here]

Columns (1) and (2) use the level of men's income (0.822 and 0.935, respectively) in 1*k* CNY/month. Columns (3) and (4) use a high-income men dummy, which represents men with incomes higher than 10*k* CNY/month. The search intensity of high-income men for beautiful women does not increase with the sex ratio in either case. High-income men are not more likely to search for a beautiful girlfriend/wife when they have more potential competition from other men. Column (2) displays this by interacting the sex ratio and the level of men's income. Column (4) shows this by interacting the sex ratio and the high-income men dummy. Column (4) does, however, show that low-income men are weakly less likely to visit beautiful women where the sex ratio is higher (-8.969).

Observation 5. Men's probability of visits to more beautiful female profiles increases with men's income, but not with the sex ratio.

Hence, although richer men search more vigorously for beautiful women, that greater search intensity does not increase with the availability of men. This result suggests high-income men are not more desperate in the face of what *should be* greater competition.

IV. Results from Census and Household Survey Data

Marriage Probability

Here we test for the effects that we have found thus far of the accumulating entry of beautiful low-income women into the mating market for high-income men on the marriage probability of high-income women. These results predict that the marriage probability of high-income women, particularly the plain-looking among them, decreases with this competitive entry, whereas that of low-income women, including the plain-looking, is not adversely affected. We use the 20 percent random sample of the 2005 China mini-Census.[21] The entire sample contains 2,585,481 individuals in 31 provinces in China.[22] Again, we

[21] 2005 is the latest year available which contains individual income data.

[22] Hong Kong, Macau, and Taiwan were excluded.

restrict the sample to women between the ages of 20 and 30 years old. Males earn a positive income, and both males and females have urban *hukou*.[23] We also excluded cities for which this mini-Census sample population of men and women in the prime age for the marriage market is below 300. Excluding the smaller cities makes our sample of cities here makes them more similar to our sample of large cities in the online dating part of our study. Our findings are only slightly less significant if we exclude cities with less than 250 of the men and women of interest. These results are available on request. Excluding cities in provinces with significant minority populations, which can exhibit unique marriage matching traditions, we obtain a final sample of 20,929 women.[24]

We estimate the following logit model of the probability of being married for woman *i*

$$P(married_i|X) = \frac{\exp(X\beta)}{1+\exp(X\beta)} \qquad \text{Eq.(2)}$$

where the dependent variable is the marital status of female *i* in city *c*. It equals 1 if the woman is married and 0 if she is single. The local sex ratio is again the *log* of the number of males 22 and 32 years old over the number of females between the ages of 20 and 30 years old in each city. The mean incomes of *H*-, *M*-, and *L*-men are defined as the top-, middle- and bottom-thirds, respectively, of the income distribution of the male populations of each city, respectively. With regard to these income levels, the average bounds across cities for men are 1,211-5,978 CNY/month for *H*-men, 757-1,123 CNY/month for *M*-men, and 194-702 CNY/month for *L*-men. The average bounds across cities for women are 1,019-3,197 CNY/month for *h*-women, 610-934 CNY/month for *m*-women, and 191-547 CNY/month for *l*-women. These ranges are not necessarily contiguous because the average bounds across cities are not the averages of bounds defined within each city.[25] We interact the dummy variables for the different categories of women with the sex ratio and the mean income of *H*-, *M*-, and *L*-men. We use the mean income of men with different income categories within a city as the treatment variable because these are exogenous to women's individual income. Again, we omit discussion of the *M*-men because the effect of the change in their mean income on high- and low-income women will be ambiguous in our theoretical framework. We also omit *L*-men because less than 6 percent of *h*-women marry them. The regression

[23] An internal passport system from the command economy era: a *hukou* entitles holders to local benefits and to government social services.

[24] The provinces we dropped are Gansu, Guangxi, Guizhou, Inner Mongolia, Ningxia, Qinghai, Tibet, Xinjiang, and Yunnan.

[25] Note also that although women earn a lower income, the average bounds of incomes of the men and the women overlap for each income category. In particular, the average lower bound of the income of low-income men (194-702 CNY/month) is not higher than the average upper bound of the income of high-income women (1,019-3,197 CNY/month).

results are presented in Table 4. Columns (1)-(3) present results for the same cities as in Table 2 for the online dating experiment. Columns (4)-(6) extend the sample to 57 cities for which we have a sample of more than 300 survey responses from each of men and women.

[Insert Table 4 here]

The weakly positive coefficient for the sex ratio in columns (1)-(6) of Table 4 for the benchmark *l*-women is consistent with our hypothesis that they may merely substitute towards high-income when more of them are available. Controlling for educational attainment and the interaction of the mean income of men with different income levels with women's income type in columns (2), (3), (5), and (6), high-income women are at least weakly more likely to be married than low-income women. Consistent with our findings of the competitive entry of beautiful low-income women into the market for high-income men when the sex ratio increases in columns (2)-(4) in Table 2, the availability of men negatively affects the marriage probability of high-income women (*sex ratio*h-women dummy*) in all specifications in Table 4. In the case of column (1), where we restrict the sample to the 15 cities of the experiment and do not introduce interactions between the mean incomes of men with different income types, the larger standard error (2.362) makes the effect of the sex ratio insignificant. However, significance returns in column (2) when we introduce the interactions and in column (4) when we increase the sample to those 57 cities with more than 300 responses from each of men and women. We use the coefficient for sex ratio (-1.938) in column (5), where we also control for the interaction of the mean income of men with different income types of women, to calculate the marginal effect of the sex ratio evaluated at the mean values of all variables. A 10 percent increase in the sex ratio decreases the probability of high-income women marrying by 4.8 percentage points compared with low-income women and by 1.2 percentage points in absolute terms, although the latter estimate suffers from a standard error that is too large to be significant.[26]

Moreover, consistent with our finding in columns (6) and (7) of Table 2 of differentiation by beauty rank among low-income women when the income of high-income men income increases, columns (2), (3), (5), and (6) in Table 4 show that the marriage probability of low-income women increases significantly with increases in the mean income of high-income men. This increase in high-income men's income should increase low-income women's

[26] Our results with married couples based on the income levels of women can be biased by their decision to participate in the labour market. We include only employed wives in our Census data. However, we do not know if these wives had reduced or planned to reduce their labour market participation prior to marriage. Nevertheless, we find similar qualitative results when we impute the wages for women using their age, educational attainment, and the number and gender of their children based on the methodology in Zhang and Liu (2003). (These results are available upon request.)

probability of marriage because it strictly improves the low-income women's options among high-income men. Consistent with this competitive entry hypothesis, the marriage probability of high-income women decreases relative to low-income women and even absolutely with respect to a zero benchmark with increases in *H*-men's mean income.

Women's marriage probability can decrease with their level of educational attainment in the West (Isen & Stevenson, 2010) and in China. This is consistent with the possibility that women who have lower marriage market endowments (e.g., attractiveness to men) have better labour market endowments or work harder (Boulier & Rosenzweig, 1984). However, this pattern is also consistent with our hypothesis that women's probability of marriage decreases with their opportunity costs, which may have a purely educational component. High-income women's probability of marriage may decrease because the number of highly educated women rises faster than the number of highly educated men, rather than as a result of the increase in competition from low-income women. To control for this possibility, columns (3) and (6) additionally control for the effect of the relative supply of men with a college or above education to women with a college or above education (*Edu ratio*). The coefficient for *Edu ratio* and *Edu ratio*h-women dummy* is not significant for column (3) but is significant for column (6), which suggests that high-income women may benefit from more educated men. The interaction between sex ratio and the high-income women dummy and the interaction between the mean income of high-income men and the high-income women dummy are only slightly changed.

We also find that women's probability of marriage decreases with the increase in dispersion in men's income, which is consistent with women waiting longer when the inequality of men increases (Gould & Paserman, 2003). However, our other coefficients are unchanged in terms of significance. (These results are available on request.)

We use the most conservative coefficient (-1.904) of the interaction between the mean income of *H*-men and *h*-women dummy in column (6) to calculate the marginal effect of the sex ratio evaluated at the mean values of all variables. A 10 percent increase in the mean income of high-income men decreases the probability of marriage for high-income women by 5.4 percentage points compared with low-income women. Women's RDP for a mate with high-income predicts this relative negative effect. When the competition for high-income men escalates, high-income women, unlike low-income women, are less disposed to substitute towards low-income men to avoid this competition.

However, this negative *relative* effect of *H*-men's income on *h*-women's probability of marriage (relative to low-income women) is also consistent with a positive *total* effect of

H-men's income on *h*-women's probability of marriage. In that case, the marriage probability of women of all income levels increases, but that of high-income women increases less than that of low-income women. Such a positive total effect is also consistent with a possible men's RDP for lower income mates. In the case of men's RDP, we expect that the first-order effect of an increase in the mean income of high-income men is to increase high-income women's marriage probability, because more of these women are of lower income than the high-income men. This is what we find in the 15 cities of the experiment. However, in the larger sample of 57 cities, we find a weakly *negative* total effect of men's mean income on high-income women's marriage probability, which is the sum of the interaction and the level of the *mean income of H-men* (-1.904+1.740=-0.164). Hence, our evidence does not support the men's RDP for lower income women hypothesis as the driver of our findings.[27]

The finding that the probability of marriage of high-income women does not increase with increases with men's incomes is remarkable because it does not support an important intuition and an empirical observation of positive assortative matching. When men are richer, more high-income women can match positively with them. However, this intuition/observation for women on average disregards the effect of increased competition from women's RDP for mate income. Further corroboration of women's RDP comes from the fact that the marriage probability of high-income women is also insignificantly affected by the incomes of low-income men (available on request). We find none of these results when we restrict the women to those above the age of 31 (available on request).

We summarize our findings with the Census data as follows.

Observation 6. The marriage probability of high-income women decreases significantly with the local sex ratio and the incomes of high-income men relative to low-income women, whereas that of low-income women increases weakly on the local sex ratio and significantly on the income of high-income men.

Somewhat supportive of standard theory, the average effect of the sex ratio on marriage probability at the bottom of Table 4 is positive for the larger sample of 57 cities with sufficient sample size, whereas it has a negative effect for the 15 cities of the online dating experiment. However, in both cases, the standard errors are too large for statistical

---

[27] Our finding that high-income women's probability of marriage decreases weakly with high-income men's income is consistent with decreasing marginal utility of income relative to the marginal utility of beauty as these men's income increases. We provide evidence for this in Observation 5. In that case, high-income men will prefer beautiful low-income women as their income increases. However, whereas men's relative decreasing marginal utility of mate income compared with mate beauty may explain why plain-looking low-income women decrease their search intensity for high-income men when these men's income increases, but it does not explain why plain-looking high-income women do not decrease their search intensity for the same men.

significance. Hence, we find some support for the standard theory that, on average, women benefit somewhat from higher sex ratio. Our results suggest, however, that the weak effect of the sex ratio on the average woman belies its strong redistributive effects across women with different income levels.

We can gain further insight into which high-income women are losing ground to the competitive entry of beautiful low-income women with the China Family Panel Studies (CFPS) 2010 baseline dataset, which has beauty ratings of surveyed subjects.[28] The CFPS is a comprehensive survey of individual-, family-, and community-level data across China, covering various aspects of economic and non-economic issues. It includes 16,000 households in 25 provinces and is representative of the whole population of China. We restrict the sample to married couples living in urban areas with a local *hukou*. We dropped the couples wherein the husband does not earn a positive income. Again, we constructed the sex ratio, which equals to the number of males between the ages 22 and 32 years of age over females between the ages of 20 and 30 years of age using data from the 2010 Census.[29] We restrict women to those between the ages of 20 and 30 years old, thereby leaving us a sample of size of 965. We use surveyor's 0 to 7 scale rating of the beauty of those they surveyed.[30]

We define the beautiful dummy = 1 for the top-20 percent of all women (rated 7 on the 1-7 scale). We again estimate the logit model of the probability of being married for woman *i* in Equation (2). Again, we define the mean income of *h*-, *m*-, and *l*-*wo*men to be the top-, middle- and bottom-third, respectively, of the income distribution of the female populations of each city. Table 5 shows the logit regression for the marriage probability for women with different income and beauty levels interacted with the local sex ratio and men's mean income in a city.

[Insert Table 5 here]

Table 5 shows that sex ratio has no effect on the marriage probability of either the plain-looking low-income women benchmark (0.945) or the beautiful-looking low-income

---

28 Although the CFPS 2012 and 2014 datasets are available, the CFPS 2010 is the last year that contains individual-level income for self-employed families.

29 We find quantitatively and qualitatively similar results when we restrict the sex ratio to include men and women between the ages of 20 and 29 years old, and test for the probability of women between the ages of 20 and 30 years old married to men to between the ages of 20 and 30 years old. We also find similar results when we use this same sex ratio and test for the probability of women between the ages of 20 and 30 years old married to men between the ages of 20 and 32 years old. The same is true when we restrict the sex ratio to include men and women between the ages of 20 and 34 years old and test for the probability of women between the ages of 20 and 30 years old married to men between the ages of 20 and 30 years old.

30 We find that the means and standard deviations of women's facial beauty do not vary systematically with the sex ratio of the provinces used in our analysis for any of the high-, medium-, and low-income levels of women. The details are available upon request. There were 269 surveyors for our sample of 965 surveyed women between the ages of 20 and 30 years old. Surveyor fixed effects are infeasible for us to introduce because raters rated on average only four subjects, which is far below the number necessary for demeaning. In many cases, there was a unique surveyor for a surveyed subject.

women (-1.238). Table 5 also shows, however, that sex ratio does exert a significantly negative effect on plain-looking high-income women (-7.294), but a weakly positive effect on beautiful high-income women (9.817). In terms of marginal effects, the marriage probability of plain-looking high-income women is 10 percentage points lower than the plain-looking low-income women and actually declines 8.7 percentage points in absolute terms when the local sex ratio increases by 10 percent. This result suggests that the decline in the high-income women's probability of marriage as the sex ratio increases observed in the interaction of the sex ratio and the *h*-women dummy in all columns of Table 4 may be limited to the plain-looking among high-income women. Unfortunately, the small sample sizes we have for many provinces (e.g., less than half of the 29 have more than 30 data points for men's income) do not permit us to study the potential linear relationship between women's marriage probability and the interaction of the mean income of men with different income types with CFPS data, as we did in Table 4 with the Census data.[31]

Observation 7. The marriage probability of the plain-looking high-income women decreases relative to plain-looking low-income women and even in absolute terms when the sex ratio increases.

We checked whether the beauty of the wife of high-income men also increases with the sex ratio. As might be expected from the increased entry of beautiful low-income women into the market for high-income men, we find that the wife of the high-income man is more attractive than that of the low-income man. Importantly, consistent with the predictable consequences of our result that a greater share of beautiful women competes for high-income men when these men become more plentiful, the beauty of the wife of high-income men increases with the sex ratio. Observation 3 suggests that the increase in the beauty of the wife of high-income men with the sex ratio is not due to the demand-side (high-income men's increased search effort for a beautiful wife). These findings (available on request) provide further evidence of our competitive entry hypothesis derived from our analysis of online dating data.

[31] We have 1114 men and 965 women in 25 provinces in the sample, with an average of 40 subjects of each gender in each province, with an actual range from 7 to 213 men and 5 to 185 for women. Since we have three types of men (by income) and six types of women (by income and beauty), there is not enough variation within some provinces with few observations. For instance, there are 9 provinces without any beautiful high-income women by our definition. If we restrict the sample to provinces with a reasonable number of observations, our regression becomes under-identified.

V. Conclusions

We use variations in men's incomes and the local sex ratio to explore the increasing burdens on high-income women from the competitive entry of low-income women into the market for high-income men due to women's RDP. When the local sex ratio or the income of high-income men increases, such that there are more high-income men or high-income men are richer, there is an increase in the search intensity of beautiful low-income women and that of the high-income women—even the plain-looking among them—for high-income men (Observation 3 and Observation 4). In contrast, only plain-looking low-income women decrease their search intensity for high-income men, when the local sex ratio or the income of high-income men increases (Observation 3 and Observation 4).

The consequence of this competitive entry of beautiful low-income women into the mating market for high-income men is evident in the marriage probability of low-and high-income women. Despite the greater search intensity of high-income women relative to low-income women, their marriage probability decreases relative to low-income women's when there are more high-income men or when high-income men are richer (Observation 6). The marriage probability of low-income women's increases weakly on the sex ratio and significantly on high-income men's mean income.

Analysis of the CPFS data reveals that it is specifically the plain-looking among high-income women whose marriage probability decreases relative to low-income women and even in absolute terms when the sex ratio increases (Observation 7). Moreover, the beauty of the wife of high-income men increases with the sex ratio. The fact that we did not find that high-income men's search for beautiful women increasing with the sex ratio (Observation 5) is consistent with the interpretation that the decrease in high-income women's marriageability when the sex ratio increases is the consequence of low-income women searching more intensively for high-income men when the sex ratio increases.

Due to the limitation of the sample size of the CFPS data for the prime age marriage market, we cannot show with direct observations of whether the decrease in the marriage probability of high-income women relative to low-income women when high-income men's income increases is mostly driven by plain-looking high-income women. Moreover, given the limitation of our experimental design of fixing high-income men's income across cities, we can show only that the change in the search intensity of the beautiful low-income women increases weakly when the income of other high-income male profiles increase. We would expect, however, that the competitive entry of beautiful low-income women to be stronger when high-income men's income increases than when the sex ratio increases. This stronger competitive entry from the increase in high-income men's income combined with these men's increased search intensity for beautiful women (Observation 5) would predict that plain-looking high-income women would be even more disadvantaged in the marriage market when high-income men's income increases than when the sex ratio increases. Therefore, we would expect plain-looking low-income women to decrease their search intensity for high-income men, because they can advantageously substitute towards the now freed-up low-income men. By contrast, high-income women, who would not avail themselves of the low-income men option when the competition for high-income men increases because of women's RDP, can only increase their search intensity for high-income men. Thus, circumstantial evidence suggests that it is the plain-looking among high-income women rather than the beautiful who experience the relative decrease in their probability of marriage when high-income men's income increases.

We find weak support for what might be predicted based on the standard theory of the effect of the sex ratios on marriage outcomes. The average women's probability of marriage weakly increases for the 57 cities with sufficient sample sizes. However, this moderate positive effect of a greater share of men in the population belies the significant redistribution of gains among women with different income levels. Our evidence suggests that beautiful low-income women gain at the expense of the high-income women, among whom the plain-looking majority loses out even in absolute terms. We emphasize that our evidence for the decreased marriage probability of high-income women when the sex ratios or high-income men's income increases most likely reflects delayed marriage rather than no marriage given the near universality of marriage for women in China. Nonetheless, our findings based on online dating field experimental, CFPS, and Census data demonstrate the novel, and to our knowledge, unexplored comparative statics effects of women's RDP for mate income. We suggest that these comparative static effects are the consequence of women's attempt to cover their labour market opportunity cost of household specialization after marriage and childbirth (which varies with their income) with the shared income of the men whom they marry.

References

Abramitzky, Ran, Delavande, Adeline, & Vasconcelos, Luis. (2011). Marrying up: the role of sex ratio in assortative matching. *American Economic Journal: Applied Economics*, *3*(July), 124–157.

Autor, David, Dorn, David, & Hanson, Gordon. (2018). When Work Disappears: Manufacturing Decline and the Falling Marriage-Market Value of Young Men. *American Economic Review: Insights, Forthcoming*.

Bartik, Timothy J. (1991). Boon or boondoggle? The debate over state and local economic development policies. In *Who Benefits from State and Local Economic Development Policies?* (Vol. 45, pp. 1–16). Upjohn Institute for Employment Research.

Becker, Gary S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, *81*(4), 813–46.

Bertrand, Marianne, Cortes, Patricia, Olivetti, Claudia, & Pan, Jessica. (2016). Social Norms, Labor Market Opportunities, and the Marriage Market Penalty for Skilled Women. *NBER Working Paper*.

Bertrand, Marianne, Kamenica, Emir, & Pan, Jessica. (2015). Gender identity and relative income within households. *Quarterly Journal of Economics*, *130*(2), 571–614.

Blau, Francine D., & Kahn, Lawrence M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, *55*(3), 789–865.

Boulier, Bryan L., & Rosenzweig, Mark R. (1984). Schooling, Search, and Spouse Selection: Testing Economic Theories of Marriage and Household Behavior. *Journal of Political Economy*, *92*(4), 712.

Bütikofer, Aline, Jensen, Sissel, & Salvanes, Kjell G. (2018). The role of parenthood on the gender gap among top earners. *European Economic Review*, *109*, 103–123.

Caucutt, Elizabeth M., Guner, Nezih, & Knowles, John. (2002). Why do women wait? Matching, wage inequality, and the incentives for fertility delay. *Review of Economic Dynamics*.

Chade, Hector, Eeckhout, Jan, & Smith, Lones. (2017). Sorting through Search and Matching Models in Economics. *Journal of Economic Literature*, *55*(2), 493–544.

Dellavigna, Stefano, Lindner, Attila, Reizer, Balázs, & Schmieder, Johannes F. (2017). Reference-dependent job search: evidence from Hungary. *Quarterly Journal of Economics*, *132*(4), 1969–2018.

Du, Julan, Wang, Yongqin, & Zhang, Yan. (2015). Sex imbalance, marital matching and intra-household bargaining: Evidence from China. *China Economic Review*, *35*(September), 197–218.

Edlund, Lena, Li, Hongbing, Yi, Junjian, & Zhang, Junsen. (2013). Sex Ratios and Crime: Evidence from China. *Review of Economics and Statistics*, *95*(5), 1520–1534.

Feng, Shuaizhang, Hu, Yingyao, & Moffitt, Robert. (2017). Long run trends in unemployment and labor force participation in urban China. *Journal of Comparative Economics*, *45*(2), 304–324.

Fincher, Leta H. (2012). China's 'leftover'women. *The New York Times*.

Ge, Suqin, & Yang, Dennis Tao. (2014). Changes in china's wage structure. *Journal of the European Economic Association*, *12*(2), 300–336.

Goldin, Claudia. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, *104*(4), 1091–1119.

Gould, Eric D., & Paserman, M. Daniele. (2003). Waiting for Mr. Right: Rising inequality and declining marriage rates. *Journal of Urban Economics*, *53*(2), 257–281.

Hamermesh, Daniel S., & Biddle, Jeff E. (1994). Beauty and the labor market. *American Economic Review*, *84*(5), 1174–1194.

Hare, Denise. (2016). What accounts for the decline in labor force participation among married women in urban China, 1991-2011? *China Economic Review*, *38*, 251–266.

Hersch, Joni. (2013). Opting out among women with elite education. *Review of Economics of the Household*, *11*(4), 469–506.

Hitsch, Günter J., Hortaçsu, Ali, & Ariely, Dan. (2010a). Matching and sorting in online dating. *American Economic Review*, *100*(1), 130–163.

Hitsch, Günter J., Hortaçsu, Ali, & Ariely, Dan. (2010b). What makes you click?-mate preferences in online dating. *Quantitative Marketing and Economics*, *8*(4), 393–427.

Hoppe, Heidrun C., Moldovanu, Benny, & Sela, Aner. (2009). The theory of assortative matching based on costly signals. *Review of Economic Studies*, *76*(1), 253–281.

Hwang, Jisoo. (2016). Housewife," gold miss," and equal: the evolution of educated women's role in Asia and the US. *Journal of Population Economics*, *29*(2), 529–570.

Isen, Adam, & Stevenson, Betsey. (2010). Women's Education and Family Behavior: Trends in Marriage, Divorce and Fertility. *Demgraphy and Economy*, 107–140.

Jiang, Quanbao, Feldman, Marcus W., & Li, Shuzhuo. (2014). Marriage Squeeze, Never-Married Proportion, and Mean Age at First Marriage in China. *Population Research and Policy Review*, *33*, 189–204.

Kleven, Henrik Jacobsen, Landais, Camille, & Søgaard, Jakob Egholt. (2018). Children and

Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics, Forthcoming*.

Knight, John, Deng, Quheng, & Li, Shi. (2017). China's expansion of higher education: The labour market consequences of a supply shock. *China Economic Review*, *43*, 127–141.

Kőszegi, Botond, & Rabin, Matthew. (2012). Reference-Dependent Consumption Plans. *American Economic Review*, *99*(3), 909–936.

Lundberg, Shelly, & Rose, Elaina. (2000). Parenthood and the Earnings of Married Men and Women. *Labour Economics*, *7*(June), 689–710.

Lundborg, Petter, Plug, Erik, & Rasmussen, Astrid Würtz. (2017). Can Women Have Children and a Career? *American Economic Review*, *107*(6), 1611–1637.

Morgan, John, Ong, David, & Zhong, Zemin (Zachary). (2018). Location still matters: Evidence from an online shopping field experiment. *Journal of Economic Behavior and Organization*, *146*, 43–54.

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

Parker, Kim, & Wang, Wendy. (2013). *Modern Parenthood: Roles of Moms and Dads Converge as They Balance Work and Family*. Pew Research Center.

Qian, Yue, & Qian, Zhenchao. (2014). The gender divide in urban China: Singlehood and assortative mating by age and education. *Demographic Research*, *31*, 1337–1364.

Rhodes, Gillian. (2006). The Evolutionary Psychology of Facial Beauty. *Annual Review of Psychology*, *57*(1), 199–226.

Shenhav, Na'ama. (2018). Lowering Standards to Wed? Spouse Quality, Marriage, and Labor Market Responses to the Gender Wage Gap. *Working Paper*.

Wei, Shang-Jin, & Zhang, Xiaobo. (2011). The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China. *Journal of Political Economy*, *119*(3), 511–564.

Weiss, Yoram, Yi, Junjian, & Zhang, Junsen. (2017). Hypergamy, Cross-boundary Marriages, and Family Behavior. *International Economic Review*, *59*(May), 757–784.

Wiswall, Matthew, & Zafar, Basit. (2016). Human Capital Investments and Expectations about Career and Family. *NBER Working Papers Series*.

Yuan, Yan, Rong, Zhao, & Xu, Lihe. (2012). Sex Imbalance, Marriage Competition, and Entrepreneurship: Evidence from rural China. *Working Paper*.

Zhang, Junsen, & Liu, Pak-wai. (2003). Testing Becker's Prediction on Assortative Mating on Spouses' Wages. *Journal of Human Resources*, *38*(1), 99–110.

Tables and Figures

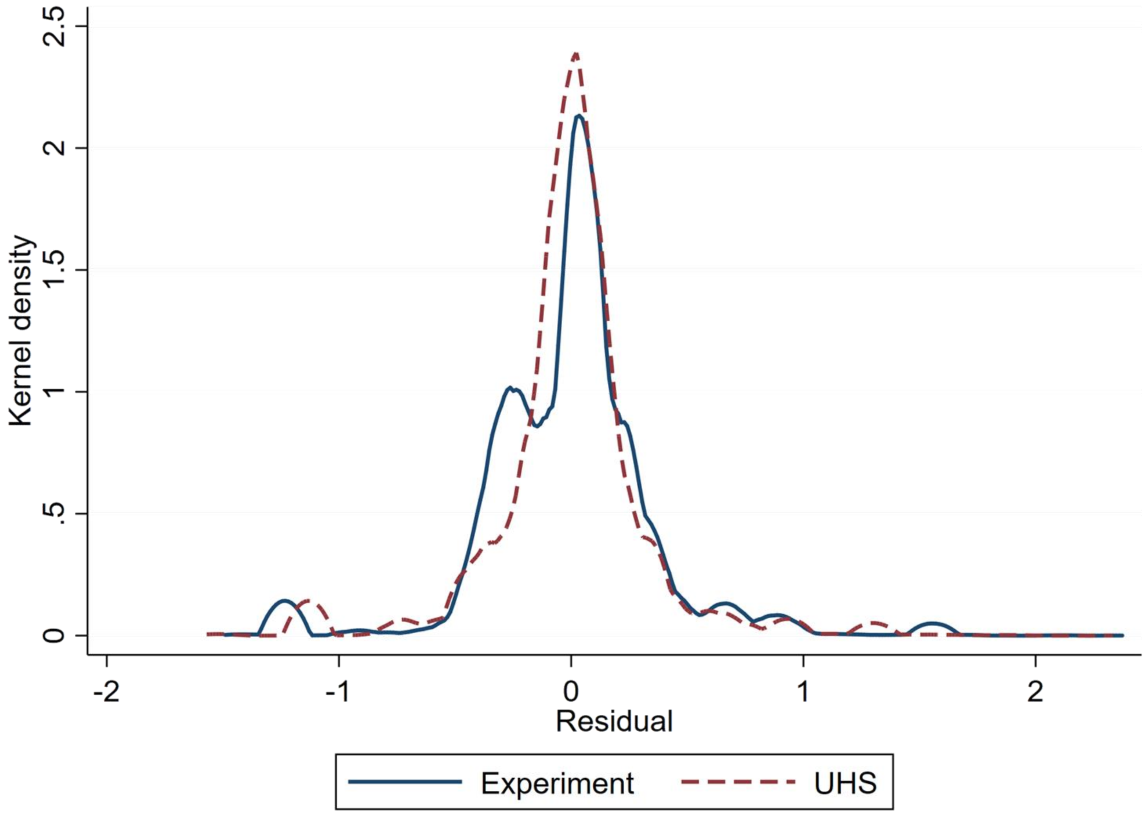


Figure 1: Residual Distribution of Women's Visitor and Single UHS Women's Income

Notes: The Urban Household Survey (UHS) from 2012 is the closest year to the year of our experiment, conducted in 2014. The cities of the UHS overlaps with the four cities used for the experiment, namely, Dalian, Shanghai, Guangzhou, and Chengdu. The residual distribution was derived from a Mincer-type regression of income on age and education for each of the UHS and the experimental datasets.

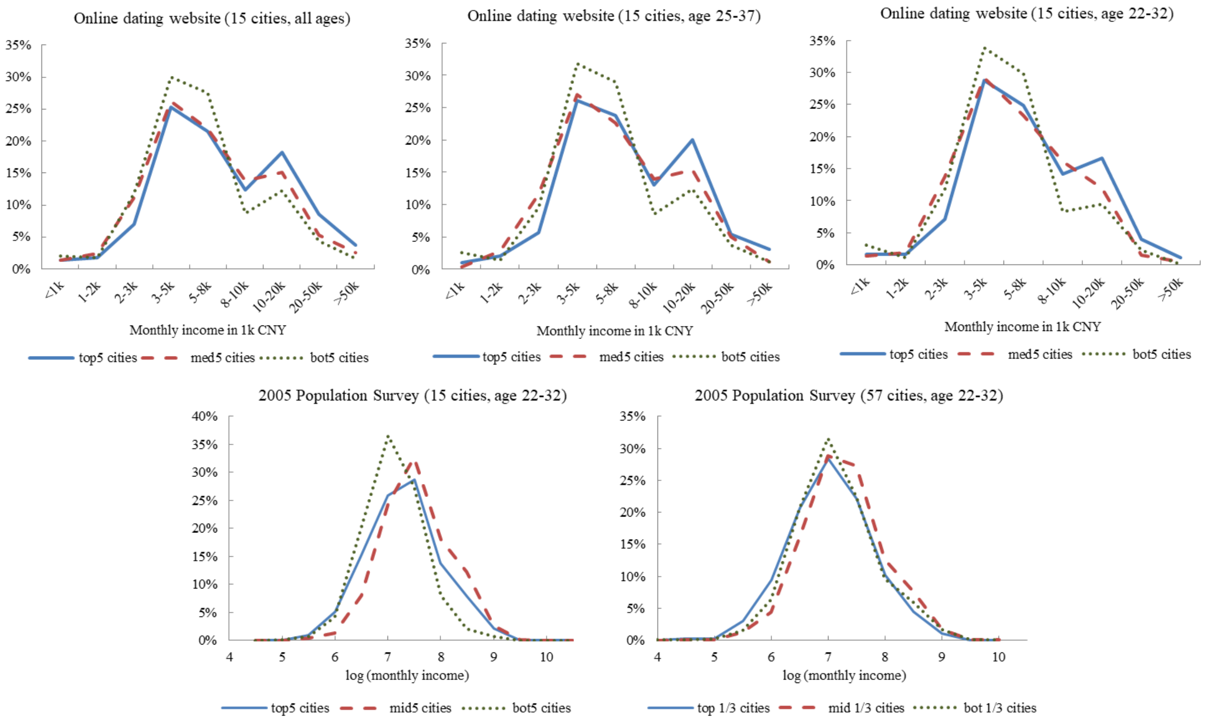


Figure 2: Men's Income Distributions on Online Dating Website and Census Data

Notes: The upper panels display the distribution of men's income on the website. The lower panel displays that for the surrounding city for the 15 cities used in the online dating experiment and the 57 cities used in our analysis of the 2005 China mini-Census data with more than 300 respondents of either gender. Within each panel, the cities are divided into top-five (top5 cities), middle- five (mid5 cities) and bottom-five (bot5 cities) five-city groups in terms of the magnitude of the local sex ratio. For the upper panels, these sex ratios are defined as the number of males between the ages of 22 and 32 years old over the number of females between the ages of 20 and 30 years old at the time of the experiments in 2014, based on the 2010 Census. The sex ratios in the lower panels are similarly defined but use the data from the 2005 mini-Census.

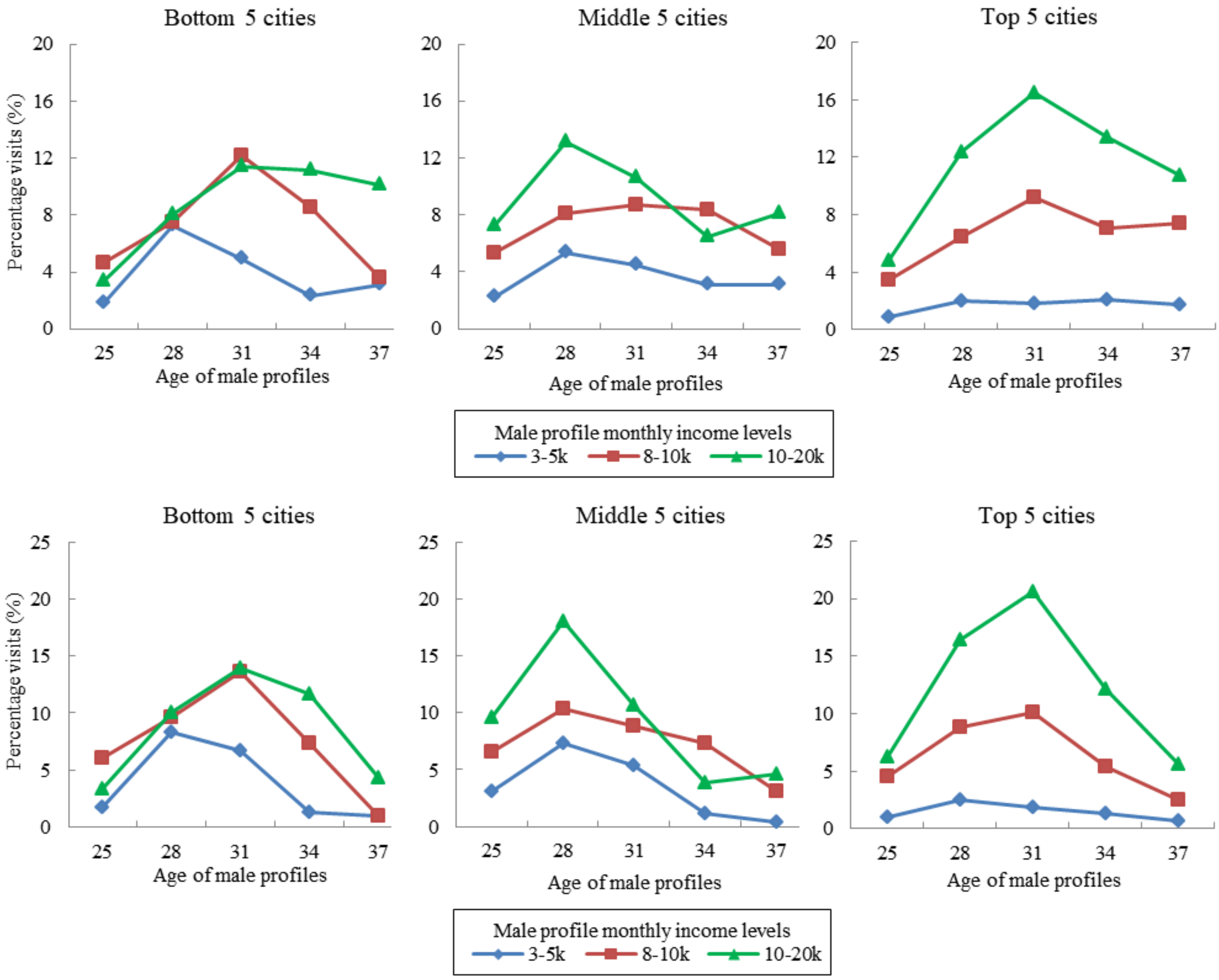


Figure 3: Share of Women's Visits by Male Profile Income Level Per Five-City Sex Ratio Group

Notes: The top panels represent the entire sample of women visitors to our profiles. They are between the ages of 18 and 45 years old. The bottom panels represent the subsample of women visitors between the ages of 20 and 30 years old. The horizontal axis in each graph indicates the age of our male profiles. The vertical axis in each graph displays the percentage of visits (%), which is the total number of visits received by each type (age and income) of male profiles in each of the five-city group divided by the visits to all our profiles for all male income types in the same five-city group. To calculate the percentage of visits, we firstly count the total number of visits received by each type (age and income) of male profiles in the top- (left-most graph), middle- (middle-graph) and bottom-five sex ratios (right-most graph) cities. This number is then normalized by the total number of visits received by all our profiles in their respective five-city group. The three lines represent the three income levels of our male profiles. Each point within each of the lines within each graph then represents the percentage of visits received by a certain type of profile among all types of profiles within that five-city group. The bottom right-most graph shows a large gap between the visit rates for men who report an income 10-20k, which peaks at 22%, and those who report 8-10k, which peaks at approximately10%, for men of all age groups. This gap is considerably lower in the left-most graph where the men who report an income of 10-20k share the same peak as those who report an income of 8-10k.

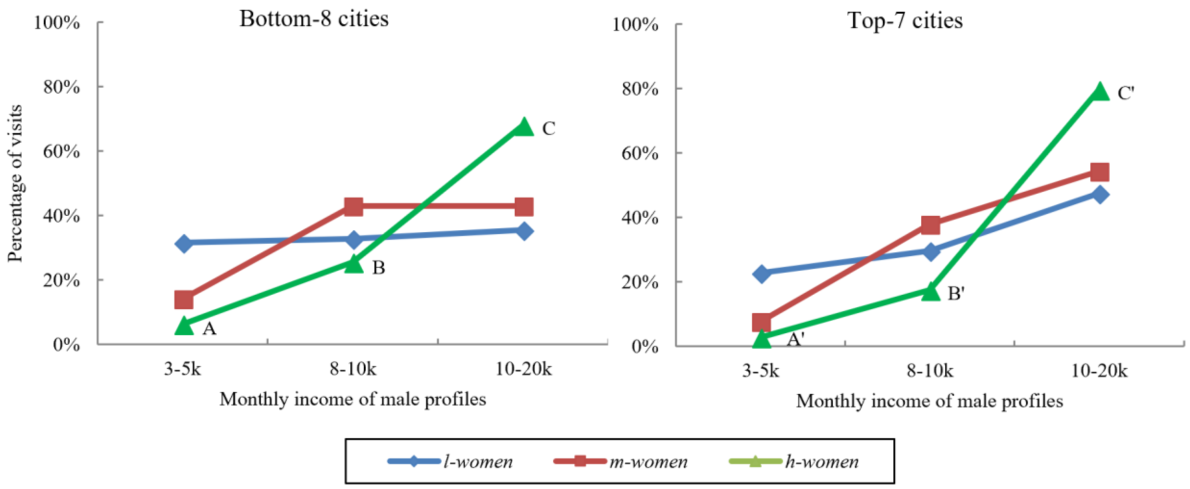


Figure 4: Share of Women's Visits to Male Profiles by Women Visitor's Income and Sex Ratio

Notes: We group women's visits into three income levels: <3, 3-5, and >5 (in 1*k* CNY), and label them as *l*-, *m*-, and *h*-women, respectively. The three lines represent the three groups of visitors. We calculate the percentage of visits of each type of women to each type of male profile. For example, on the left side, the percentage of visits of *h*-income women to high-income (10-20k) men is approximately 70 percent, in contrast to their visits in top-seven sex ratio cities, where it is 80 percent. All three points in each line add up to 100%. The lines for the top-seven sex ratio cities are rotated versions for those of the bottom-eight, indicating that women visited our high-income profiles more than our low-income profiles in the top-seven cities.

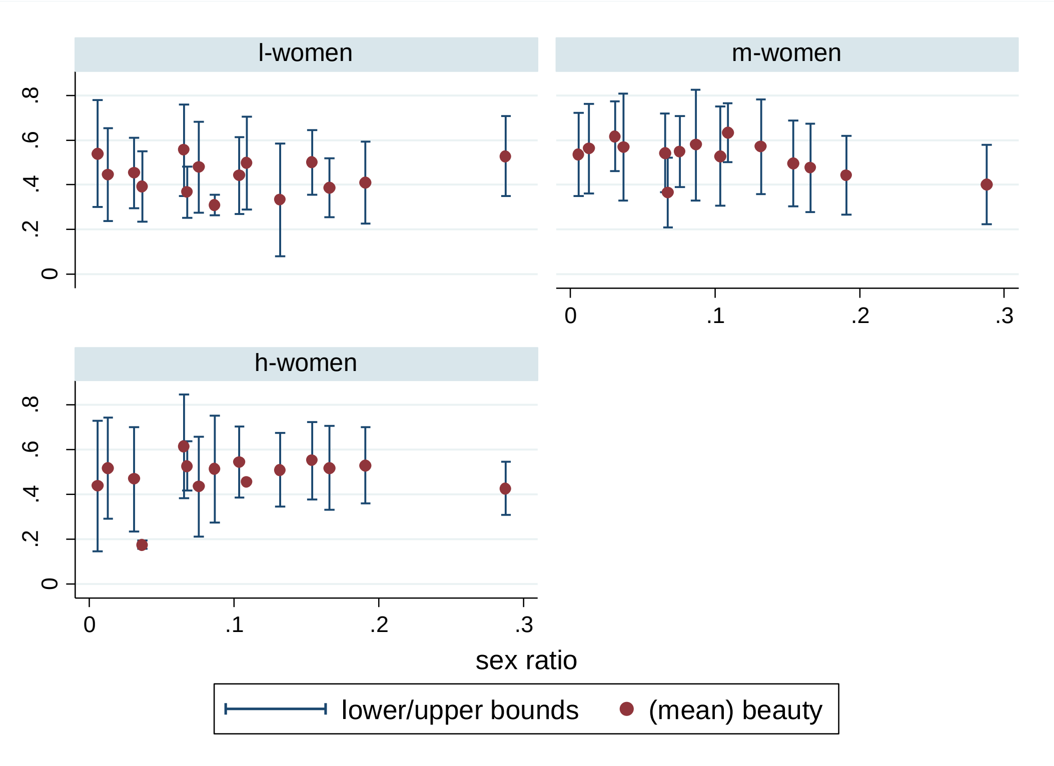


Figure 5: Distribution of Beauty of Women Visitors by Income Level vs. Log Sex Ratio

Notes: The upper and lower bounds are one standard deviation above and below the mean, respectively. The women's visits are grouped into three income levels: <3, 3-5, and >5 (in 1*k* CNY), which are labelled as *l*-, *m*-, *and h*-women, respectively. Within each panel, the vertical axis represents the woman visitor's profile picture beauty rank (0-1). The horizontal axis in each panel represents the log sex ratio of the city of the woman visitor.

Table 1: Regression Comparing Women's Visitor Characteristics and UHS Sample

Base group: UHS sample	Log income		
	Age 18-45	Age 20-30	Age 20-30
	(1)	(2)	(3)
Age	0.259**	0.964	0.061**
	(0.071)	(0.445)	(0.015)
Age^2	-0.004**	-0.018	
	(0.001)	(0.009)	
Edu years	0.095**	0.065	0.090
	(0.028)	(0.048)	(0.043)
Experiment sample	1.178	7.266	1.034
	(1.939)	(7.679)	(0.899)
Age*Experiment sample	**-0.055**	**-0.543**	**-0.013**
	(0.154)	**(0.622)**	**(0.031)**
Age^2*Experiment sample	**0.001**	**0.011**	
	(0.002)	**(0.012)**	
Edu years*Experiment sample	**0.026**	**0.012**	**-0.014**
	(0.039)	**(0.046)**	**(0.044)**
City FE	Y		
Constant	1.967	-6.281	4.685***
	(0.865)	(5.146)	(0.302)
Observations	12,758	10,545	10,545
R-squared	0.551	0.598	0.510

Notes: The Urban Household Survey (UHS) from 2012 is the closest year to the year of our experiment, conducted in 2014. The cities of the UHS overlap with four of the cities used in our experiment, namely, Dalian, Shanghai, Guangzhou, and Chengdu. The sample size for the experiment is 435 observations, whereas that for the UHS data is 12,323 observations. FE means fixed effects. Robust standard errors clustered at the city-level are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

Table 2: Ordered Probit Regression of Women's Visits on Male Profile Income

Dependent variable:	Income level of male profile visited (low (3-5*k*), middle (8-10*k*), high (10-20*k*))							
	Age 18-45	Age 18-45	Age 20-30	Age 20-30 (2SLS)	Age 31-45	Age 20-30	Age 20-30 (2SLS)	Age 31-45
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
m-women dummy	0.156	-0.636	-0.993*	-0.889	-0.210	-1.619	-1.027	-1.464
	(0.113)	(0.508)	(0.543)	(0.567)	(0.788)	(1.172)	(1.265)	(1.426)
h-women dummy	0.502***	-0.736*	-1.727***	-1.943***	0.117	0.006	-0.308	-0.487
	(0.159)	(0.393)	(0.620)	(0.714)	(0.754)	(1.595)	(1.428)	(1.232)
Sex ratio	**0.432**	**-6.219****	**-11.381*****	**-6.512**	**5.093**	**-4.289**	**3.793**	**-3.215**
	(0.638)	(2.491)	(3.485)	(5.336)	(3.577)	(3.532)	(6.261)	(3.149)
Sex ratio*m-women dummy	**0.891**	**5.401***	**10.285****	**8.854**	**-3.895**	**0.804**	**0.241**	**4.315*****
	(0.859)	(3.222)	(5.157)	(5.469)	(6.485)	(1.721)	(1.597)	(1.620)
Sex ratio*h-women dummy	**1.765****	**10.879*****	**15.861*****	**15.826*****	**5.657**	**-1.962**	**-2.599**	**7.053****
	(0.892)	(2.678)	(4.770)	(5.773)	(5.285)	(1.953)	(1.665)	(2.954)
Beauty		-1.184	-1.973**	-1.142	2.006**	-3.494*	-3.099	0.853
		(0.800)	(0.821)	(0.980)	(0.905)	(1.929)	(2.036)	(1.292)
Beauty**m*-women dummy		1.172	1.713*	1.466	-0.133	4.001*	2.888	4.565*
		(0.967)	(1.008)	(1.042)	(1.634)	(2.263)	(2.359)	(2.603)
Beauty**h*-women dummy		1.957***	3.722***	3.960***	-1.321	2.105	2.389	-0.497
		(0.736)	(1.124)	(1.284)	(1.931)	(2.913)	(2.653)	(2.237)
Sex ratio*beauty		**12.131*****	**20.817*****	**13.027**	**-20.534*****	**9.518***	**-3.466**	**-2.816**
		(4.277)	(5.097)	(8.138)	(6.764)	(5.680)	(10.222)	(6.693)
Sex ratio*beauty*m-women dummy		**-8.339**	**-16.738***	**-13.786**	**14.794**			
		(7.248)	(9.311)	(9.705)	(16.266)			
Sex ratio*beauty*h-women dummy		**-19.219*****	**-29.824*****	**-28.938*****	**7.079**			
		(4.912)	(7.906)	(9.855)	(12.547)			
Mean income H-men						**-0.483****	**-0.649*****	**0.348**
						(0.201)	(0.208)	(0.237)
Mean income H-men*beauty						**0.567**	**0.868****	**-0.284**
						(0.402)	(0.403)	(0.495)
Mean income H-men*m-women dummy						**0.615*****	**0.674*****	**-0.137**
						(0.187)	(0.207)	(0.265)
Mean income H-men*beauty*m-women dummy						**-0.853****	**-1.001*****	**-0.426**
						(0.368)	(0.375)	(0.563)
Mean income H-men*h-women dummy						**0.990*****	**0.927*****	**-0.276**
						(0.288)	(0.279)	(0.260)
Mean income H-men*beauty*h-women dummy						**-1.370****	**-1.322*****	**0.084**
						(0.538)	(0.499)	(0.657)
Additional controls:								
Age and education dummies of female visitors	Y	Y	Y	Y	Y	Y	Y	Y
Mean and standard deviation of men's and women's incomes in each city		Y	Y	Y	Y	Y	Y	Y
Observations	1,760	867	548	548	308	548	548	308
Pseudo R^2	0.049	0.066	0.085	0.085	0.088	0.098	0.098	0.117

Notes: Data from the online dating experiment, wherein each observation is a visit (click) from a woman visitor. The local sex ratio is defined as the number of males between the ages of 22 and 32 years old over females between the ages of 20 and 30 years old at the time of experiment in 2014, proxied by males between the ages of 18 and 28 years old and females between the ages of 16 and 26 years old in the 2010 Census. *h*-, *m*-, and *l*-men indicate top-, middle- and bottom-third women by monthly income in each city, respectively. The *l*-women are the omitted benchmark with income less than 3*k* CNY/month. *m*-women dummy = 1 if the woman's income is 3-5*k* CNY/month. *h*-women dummy = 1 if the woman's income is more than 5*k* CNY/month. *H*-, *M*-, and *L*-men denote top-, middle- and bottom-third of male visitors to our female profiles by monthly income in each city, respectively. The results for *M*-and *L*-men are suppressed and available on request. Beauty is the beauty percentile ranking of female visitors which we acquire for a random sample of 2/3 of visits. Columns (4) and (7) are the second stage of the Bartik-type IV regression results for columns (3) and (6), respectively. The first stage is in A-Table 8 of Appendix 3. The mean and standard deviation of men's and women's incomes are based on the online dating users and defined in 1k CNY at the city-level. Robust standard errors clustered at the city-level are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

Table 3: OLS Regression of Men's Visits on Female Profile's Beauty

Dependent variable	Beauty ranking (0-100) of female profile visited			
	(1)	(2)	(3)	(4)
Income of men	0.822***	0.935**		
	(0.267)	(0.343)		
Sex ratio		-3.131		-8.969
		(21.026)		(5.638)
Income of men*sex ratio		-0.909		
		(2.568)		
H-men dummy			1.341**	1.739**
			(0.584)	(0.790)
Sex ratio**H*-men dummy				-3.416
				(5.762)
Additional controls:				
Age and education dummies of male visitors	Y	Y	Y	Y
Mean and standard deviation of men's and women's incomes in each city	Y	Y	Y	Y
Constant	43.543*	31.034	50.095**	37.538
	(22.873)	(24.894)	(21.643)	(23.678)
Observations	5,288	5,288	5,288	5,288
R^2	0.043	0.044	0.042	0.044

Notes: Data from another experiment which was run simultaneously with 390 female profiles in the same 15 cities. The income of men is measured in 1*k* CNY/month. These profiles of females had ages of 22, 25, 28, 31, and 34, with a height of 163 cm, were college educated, and had incomes of 5-8k CNY/month. The local sex ratio is calculated in the same way as in Table 2. *L*-men is the omitted benchmark in column (3) and (4) with income less than 5*k* CNY/month. *M*-men = 1 if the men's income is between 5k and 10k CNY/month. The results for *M*-men are suppressed and available on request. *H*-men = 1 if men's income is more than 10k CNY/month. Robust standard errors clustered at the city-level are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

Table 4: Logit Regression of Women's Marriage Probability (Census data)

Dependent variable	Logit: 1 = married, 0 = single					
	15 cities in experiment			57 cities		
	(1)	(2)	(3)	(4)	(5)	(6)
Sex ratio	0.120	2.270	2.450	1.520	1.425	1.652*
	(2.472)	(2.045)	(1.686)	(0.964)	(0.967)	(0.988)
h-women dummy	-0.210	0.127	0.389	-0.110	1.109	0.927
	(0.180)	(2.396)	(2.558)	(0.099)	(1.348)	(1.358)
Sex ratio*h-women dummy	**-2.308**	**-4.579****	**-4.452****	**-1.760****	**-1.938****	**-2.098****
	(2.362)	(1.813)	(1.834)	(0.856)	(0.935)	(0.924)
Edu ratio (men BA+/women BA+)			-0.793			-0.282
			(0.563)			(0.183)
Edu ratio**h*-women dummy			0.165			0.273*
			(0.489)			(0.157)
Mean income of *H*-men		6.179***	5.810***		1.891**	1.740*
		(1.580)	(1.584)		(0.937)	(0.892)
Mean income of H-men*h-women dummy		**-4.849*****	**-4.759*****		**-2.159*****	**-1.904*****
		(1.445)	(1.497)		(0.713)	(0.715)
Additional controls:						
Age and education dummies of women	Y	Y	Y	Y	Y	Y
Mean and standard deviation of men's and women's incomes in each city	Y	Y	Y	Y	Y	Y
Constant	0.329	-2.480	-0.506	1.089	0.710	0.927
	(2.388)	(3.442)	(1.388)	(1.184)	(1.855)	(1.893)
Observations	10,888	10,888	20,929	20,929	20,929	20,929
Pseudo R^2	0.287	0.292	0.298	0.298	0.299	0.299
	Average effect of sex ratio					
	0.120	2.270	2.450	1.520	1.425	1.652
	+(-1.778)/3	+(-3.945)/3	+(-4.064)/3	+(-0.504)/3	+(-0.354)/3	+(-0.344)/3
	+(-2.308)/3	+(-4.579)/3	+(-4.452)/3	+(-1.760)/3	+(-1.938)/3	+(-2.098)/3
	= -1.242	= -0.571	= -0.389	= 0.765	= 0.661	= 0.838

Notes: Data are from the 2005 China mini-Census, restricted to men and women with an urban *hukou*, and positive monthly income. Women are restricted to those between the ages of 20 and 30 years old. The local sex ratio is defined as the number of men between the ages of 22 and 32 years old over the number of women between the ages of 20 and 30 years old in the 2005 mini-Census in each city. *H*- (*h*-), *M*- (*m*-), and *L*-(*l*-) men (women) are defined as top-, middle- and bottom-1/3 men (women) by monthly income in each city, respectively. The *l*-income women are the omitted benchmark. Results for *M*-men and *m*-women are suppressed and available on request. *Edu ratio* is defined as the number of males with a bachelor's degree or higher over the number of females with a bachelor's degree or higher in the city. *Sex ratio*, *edu ratio*, and all incomes are in *log* form. To calculate the average effect of sex ratio, denote the coefficients for *sex ratio*, *sex ratio*m-women dummy*, *sex ratio*h-women dummy* by *a*, *b*, and *c*. The marginal effect of sex ratio on *l*-women is *a*, on *m*-women is *(a+b)*, and on *h*-women is *(a+c)*. Given that women are divided into three groups equally, the average effect is $a/3 + (a+b)/3 + (a+c)/3 = a + b/3 + c/3$. Robust standard errors clustered at the city-level are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1

Table 5: Logit Regression of Women's Probability of Marriage by Beauty (CFPS data)

Dependent variable	
	Logit: 1 = married, 0 = single
Sex ratio	0.945
	(3.173)
h-women dummy	-0.332
	(0.305)
Sex ratio*h-women dummy	**-7.294****
	(3.183)
Beautiful dummy	0.823**
	(0.370)
Beautiful dummy**h*-women dummy	-0.607
	(0.423)
Sex ratio*beautiful dummy	-1.238
	(3.756)
Sex ratio*h-women dummy*beautiful dummy	**9.817**
	(6.535)
Mean income of men	0.023
	(0.383)
Additional controls:	
Age and education dummies of women	Y
Mean and standard deviation of men's and women's incomes in each province	Y
Constant	-0.838
	(2.738)
Observations	898
Pseudo R2	0.344

Notes: Data are from China Family Panel Studies (CFPS) 2010. The CFPS provides the residency status of individuals and households only at the provincial level. The sample is restricted to men between the ages of 22-32 years old and women between the ages of 20 and 30 years old with an urban *hukou,* and positive monthly income. The local sex ratio defined as the number of males between the ages of 22 and 32 years old over the number of females between the ages of 20 and 30 years old in each province is calculated using the 2010 Census. The sex ratio and all incomes are in *log* form. *h*-, *m*-, and *l*-women are defined as top-, middle- and bottom-1/3 women by monthly income in each province, respectively. The *l*-women are the omitted benchmark. The results for *m*-women are suppressed and available on request. Beautiful dummy = 1 if a woman has a beauty rating of 7 out of 1-7 scale in the CFPS. Robust standard errors clustered at province level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendices (Not for publication)

Appendix 1. A Game Theoretic Illustration of the Competition between High- and Low-income Women[1]

A-Table 1: Game matrix for competition between high- and low-income women for high-income men

		High-income woman	
		$No\ effort\ \ (1-e)$	$Effort\ \ (e)$
Low-income woman	$Try\ \ (t)$	$\theta - c, 0$	$z\theta - c, (1-z)\theta - c$
	$Not\ Try\ (1-t)$	$1, \theta$	$1, \theta - c$

We show in this two-player contest example how high-income women can be negatively affected in the marriage market when the prize they seek exclusively (high-income men) increases in value due to the increased competition (that is, search effort/effort at attraction) from low-income women. We also show that low-income women's effort is influenced by their beauty to a higher degree than high-income women.

For ease of exposition, we make a number of simplifications. We can model the competition between high- and low-income women as the competition between two types of players: a high-income woman (Column Player) and a low-income woman (Row Player), because our focus here is not the within-income group competition among women but the across-income group competition (see the game matrix in A-Table 1). Since our purpose is to model the impact of effort on search outcomes and not search outcomes per se, we do not model the actual search process.

We model the choices of each player for 2x2 outcomes in this game: $\{(Try, Not\ Try) \times (Effort, No\ effort)\}$. The payoffs for the low-income woman are the first coordinates of each pair for each outcome as represented in the above matrix, whereas that of the high-income woman is the second. These payoffs are the values of the "prizes" of competition, which are two men: one high-income man and

[1] We are grateful to Barton Lipman for developing this example with us. All errors are ours.

one low-income man, both of whom are passive players. The low-income woman automatically "gets" the low-income man, who she values at 1, or can Try for the high-income man, who she values at θ minus her cost of effort (from searching or otherwise, e.g., putting more effort in grooming) c. The high-income woman, who also values the high-income man at θ can put in $Effort$ at cost c in getting him. She has no interest in the low-income man. Hence, the outer nest (that is, the low-income men) of the nested prize (that is, high- and low-income men) structure mentioned in the introduction is implicitly an option only for the low-income woman, the value of which is fixed at 1. Fixing the value of this option allows us to focus on how an increase in the value of the common prize (the high-income man) θ changes the competition between these two types of women.[2]

We model the comparative advantage of each type of woman in the competition for the high-income man by specifying that if the low-income woman chooses Try to get the high-income man and the high-income woman also chooses $Effort$, the low-income woman succeeds with probability z, which can be anything strictly between 0 and 1. Otherwise, if one of the two is trying (putting in effort) to get the high-income man and the other is not, the one who is trying succeeds for sure. If neither is trying, the high-income woman gets him. This last assumption is simplifying and raises the bar for our goal, which is to show that the high-income woman can be negatively affected when θ increases.

To predict what players do in equilibrium, we must find choices that are mutually enforcing, ones in which no player wants to deviate from their equilibrium strategy (a probability distribution over her respective pair of actions) given her opponent's equilibrium strategy. Let e stand for the probability high-income woman chooses $Effort$ and t stand for the probability that the low-income woman chooses Try. Assume $\theta > 1$, that is, getting the high-income man yields a higher payoff than getting the low-income man. Also, restrict $\theta - c > 0$, $z\theta - c > 0$ and $(1-z)\theta - \mathrm{c} > 0$ so

[2] Assuming that the marital surplus is largely exogenous (e.g., due to the predominant role of public goods over private goods), our model would fit within the non-transferable utility framework. Such public goods as a house, a minimal level of financial security, or children, provided by the husband is increasing with his income. In the case of a minimum level of financial security in marriage, women may seek that either for its own sake or because they plan to withdraw from the labour market after marriage and childbirth. In that case, high-income women have no greater ability to compensate their husband than low-income women because neither may participate in the labour market. The fixed cost c can be interpreted as either search cost within the context of directed search for high-income men (Chade et al., 2017) or the additional effort necessary to signal to attract high-income men (Hoppe, Moldovanu, & Sela, 2009). We thank P. A. Chiappori for helping us clarify the relation of our study to the marriage matching theory literature.

that the payoff in each case is non-negative. We first look for pure strategy Nash equilibria, which is represented as ordered pairs of actions equivalent to a pair of degenerate probabilities over those actions. We derive the mixed strategy Nash equilibria, which are represented by a pair of probabilities (t, e), each strictly between 0 and 1.

For $(Not\,try, No\,effort)$ to be equilibrium, low-income woman chooses $Not\,try$ given high-income woman chooses $No\,effort$, and high-income woman chooses $No\,effort$ given low-income woman chooses $Not\,try$. For low-income woman to choose $Not\,try$ requires $1 > \theta - c$. For high-income woman to choose $No\,effort$ requires $\theta > \theta - c$. In other words, $\theta < 1 + c$. In this pure strategy Nash equilibrium, $e = 0$ and $t = 0$. The payoff for high- and low-income woman is θ and 1, respectively.

For $(Not\,try, Effort)$ to be equilibrium, low-income woman chooses $Not\,try$ given high-income woman chooses $Effort$, and high-income woman chooses $Effort$ given low-income woman chooses $Not\,try$. The low-income woman requires $1 > z\theta - c$, whereas the high-income requires $\theta - c > \theta$, which is impossible.

For $(Try, No\,effort)$ to be equilibrium, low-income woman chooses Try given high-income woman chooses $No\,effort$, and high-income woman chooses $No\,effort$ given low-income woman chooses Try. The low-income woman requires $\theta - c > 1$, whereas the high-income woman requires $0 > (1 - z)\theta - c$, which is impossible under our restriction $(1 - z)\theta - c > 0$.

For $(Try, Effort)$ to be equilibrium, low-income woman chooses Try given high-income woman chooses $Effort$, and high-income woman chooses $Effort$ given low-income woman chooses Try. The low-income woman requires $z\theta - c > 1$, whereas the high-income woman requires $(1 - z)\theta - c > 0$. Together, we obtain $\theta > \frac{1+c}{z}$. In this pure strategy Nash equilibrium, $e = 1$ and $t = 1$. The payoff for high- and low-income woman is $(1 - z)\theta - c$ and $z\theta - c$, respectively.

Next, we look for the interior mixed strategy equilibrium. This equilibrium requires that the low-income woman is indifferent between Try and $Not\,try$, given the high-income woman's strategy, and the high-income woman is indifferent between $Effort$ and $No\,effort$, given the low-income woman's strategy. In other words, it requires

$$(1-e)(\theta-c)+\ e(z\theta-c)=1 \qquad \text{Eq.(3)}$$

$$t\big((1-z)\theta-c\big)+(1-t)(\theta-c)=t\cdot 0+(1-t)\theta \qquad \text{Eq.(4)}$$

Solving the two equations gives us $e=\frac{\theta-1-c}{(1-z)\theta}$ and $t=\frac{c}{(1-z)\theta}$. The interior mixed strategy equilibrium requires e and t to be strictly between 0 and 1, which in turn requires $1+c<\theta<\frac{1+c}{z}$. Plugging $e=\frac{\theta-1-c}{(1-z)\theta}$ and $t=\frac{c}{(1-z)\theta}$ back into the above equations, the payoff for high- and low-income woman is $\frac{(1-z)\theta-c}{1-z}$ and 1, respectively.

All results are summarized in A-Table 2, which shows that if θ is below $1+c$, then it is a dominant strategy for the low-income woman to $Not\ try$. The high-income woman's payoff is θ. When θ increases from below $1+c$ to above $1+c$, the low-income woman chooses Try with strictly positive probability. As a consequence, the high-income woman's payoff drops discontinuously from θ to $\frac{(1-z)\theta-c}{1-z}=\theta-\frac{c}{1-z}$. If we increase θ further to be above $\frac{(1+c)}{z}$, the low-income woman tries with probability 1, and high-income woman's payoff again drops discontinuously, but this time from $\frac{(1-z)\theta-c}{1-z}$ to $(1-z)\theta-c$. Although the high-income woman's payoff increases with θ for some range, crucially in support of our empirical results, it decreases discontinuously as θ goes up further, because the low-income woman chooses non-zero levels of effort. These equilibrium strategies and payoffs detailed in A-Table 2 are further illustrated in A-Figure 1.

We need only change the interpretation of this game slightly to model the effect of an increase in sex ratio. Let the two types of women now be two populations of otherwise homogenous individual women, namely high- and low-income.[3] We interpret the probability distribution of their equilibrium strategies as the share each type of women adopting these strategies. Let z represent the *share* of high-income men that the low-income women population gets given the shares of both the high- and the low-income women populations that put in $Effort$ or Try, respectively. When the sex ratio increases, the high-income men are less scarce. The *ex-ante* effect of this

---

[3] The probability with which individual players choose an action in a mixed strategy Nash equilibrium can be interpreted as shares of a population of players choosing pure strategies. See Harsanyi's purification theorem for details: https://en.wikipedia.org/wiki/Purification_theorem

decrease in the scarcity of high-income men prizes can be modelled now by substituting θ with $s \cdot \theta$, where $s \geq 0$ increases with the population of high-income men. It is obvious that we would find similar results with an increase in s that we find with an increase in θ. Hence, the effect of an increase in the sex ratio *ex-ante* to equilibrium is similar to an increase in the income of high-income men. However, the effects of an increase in the sex ratio in equilibrium, which takes into account potential shifts in competition between and within women from different income groups, are more subtle, especially if men and women also differ by other characteristics, such as beauty.

In real life, not only is the expected value of pursuing high-income men *ex-ante* to equilibrium higher when the sex ratio increases, but the probability of getting a better-looking low-income man is also higher because there are a greater number of men for every woman. Therefore, women in general not only have a better chance of getting a high-income men, but women with different levels of income and beauty have more scope to adopt heterogeneous strategies because there are more men. For example, if beautiful women of high- and low-income levels pursue high-income men more because of increased availability, plain-looking low-income women would face stiffer competition for these men. These plain-looking low-income women may rather pursue better-looking low-income men more—despite the greater availability of high-income men. Accordingly, only certain subsets of women may actually enjoy the greater availability of high-income men in equilibrium, whereas others are crowded out.

The effect of this additional level of heterogeneity based on beauty can be captured within our simple framework by the introduction of heterogeneous costs, where the cost of the plain-looking women is higher than that of the beautiful looking women: $c_p > c_b$. The threshold θ_p for the plain-looking low-income women to Try must be higher than that for beautiful-looking low-income women θ_b since their expected value of trying depends on their cost and $z\theta - c_p < z\theta - c_b$. However, plainness does not affect high-income women's propensity to choose $Effort$; they only choose effort because plain-looking low-income women are trying, since by assumption $(1 - z)\theta - c_p > 0$.

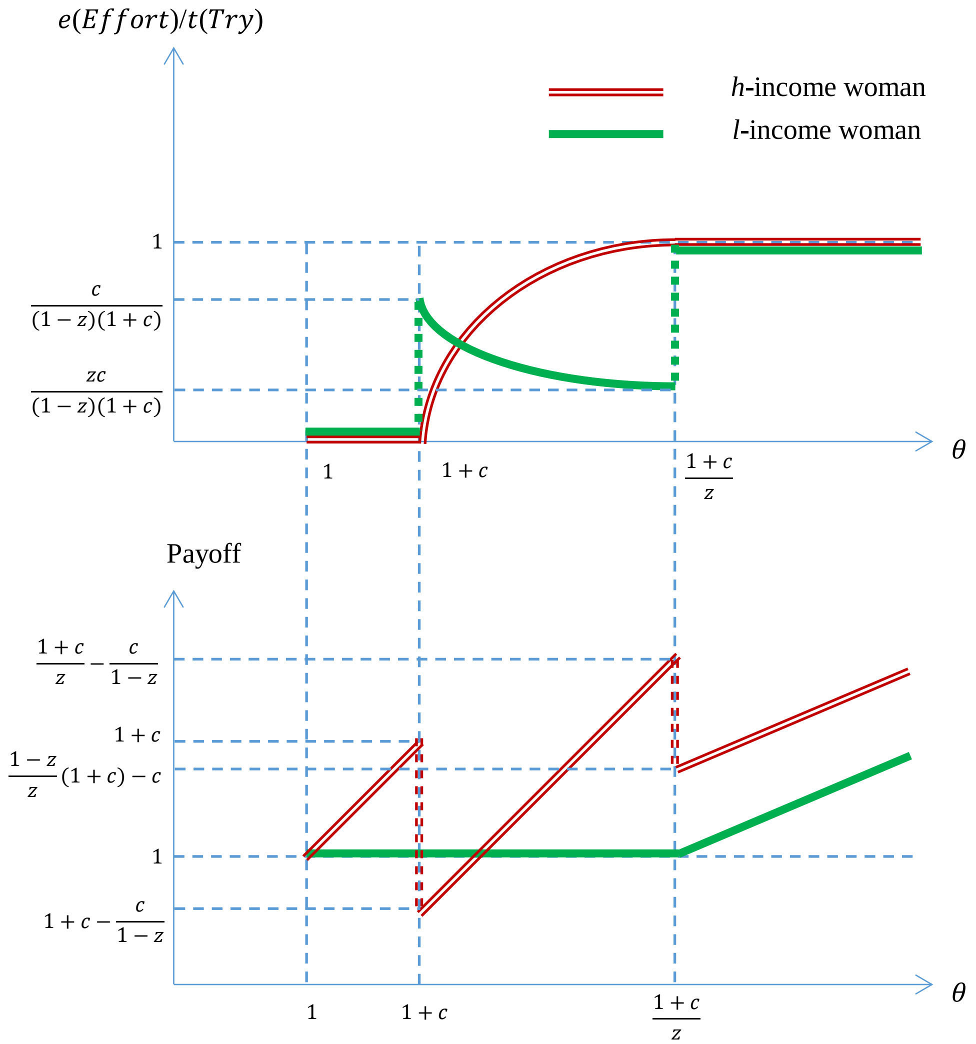
We leave the detailed modelling of these very interesting potential equilibrium effects for future research, as our focus in this paper is empirical.[4] Our main goal here is to show that high-income women can be negatively affected by the increase in the sex ratio or the increase in the income of high-income men due to the consequent increase in the entry of low-income women.

A-Table 2: Equilibrium Payoffs for Each Type of Woman Given z and θ

	$1 < \theta < 1 + c$	$1 + \mathrm{c} < \theta < \frac{(1+c)}{z}$	$\theta > \frac{(1+c)}{z}$
h-woman			
$e(Effort)$	0	$\frac{\theta-1-c}{(1-z)\theta}$	1
payoff	θ	$\frac{(1-z)\theta-c}{1-z}$	$(1-z)\theta-c$
l-woman			
$t(Try)$	0	$\frac{c}{(1-z)\theta}$	1
payoff	1	1	$z\theta-c$

Note: The top row details the probability of $Effort$ for the high (*h*)-income woman and her payoff in equilibrium for a given values of θ, and the bottom row details that of Try for the low (*l*)-income woman.

4 We present empirical evidence that lower income women adopt heterogeneous strategies according to their beauty when the sex ratio (Observation 3) and the income of high-income men (Observation 4) increase.



A-Figure 1: Strategies and Payoffs for High- and Low-income Women

Notes: θ is the value of the high-income man. The top panel illustrates the equilibrium strategy for the high (h)- and low (l)-income women, whereas the bottom panel illustrates their respective payoffs. The thin double-lines are those of the high-income woman. The thick green lines are those of the low-income woman.

An important and perhaps counterintuitive qualitative result from the top part of A-Figure 1 is that although the high-income woman is negatively affected by the entry of the low-income woman into the competition for high-income men, the low-income woman's probability of Try initially jumps above the high-income woman's increasing probability of $Effort$ at $1+c$ but then decreases and crosses the high-income woman's probability of $Effort$ to decrease to a lower level than the high-income woman's probability of $Effort$.

One way of interpreting this result is that the low-income woman will only give up her low-income man outside option from whom she gets a sure payoff of 1 to take a risk in obtaining θ or zero, if she gets the high-income man with a high enough probability. She can only ensure that her probability of winning is high enough if she chooses Try with sufficiently high probability. In contrast, the high-income woman has no such option, and hence, increases her probability of $Effort$ continuously when $\theta > 1 + c$.

The high-income woman has the advantage that she gets the high-income man by default. Moreover, she will not be challenged until the low-income woman is compensated for the loss of the low-income man and the search effort of c for the high-income man. Such challenge can be interpreted as search friction. However, as the value of θ increases further, the high-income woman's probability of $Effort$ will increase, decreasing the returns of the low-income woman in choosing Try, causing the low-income woman to decrease her probability of Try.

The expected payoff of the high-income woman, which models the probability of marriage of high-income women to the high-income men, drops when $\theta > 1 + c$, as the low-income woman enters the competition for the high-income man, and then increases linearly. In the real world, this greater number of women desiring the same men effect will likely be more continuous and predict a continuous decrease in the probability of marriage of high-income women. However, as the sex ratio increases further, the effect of a greater number of desirable men may dominate the effect of a greater number of women desiring the same men. Hence, our simple model predicts a non-monotonic effect of the increase in either sex ratio or the income of high-income men.

Appendix 2. Background on Cities and Visitors

We started with 36 major cities (including all 31 provincial capitals and five vice-provincial level cities). We excluded 10 cities in minority provinces, and Ningbo, which is very close to Shanghai and Hangzhou, and Shenzhen which is too close to Hong Kong and may be affected by the Hong Kong marriage market. We also excluded three cities between the ages of 20 and 29 years old and 25 and 34 years old sex ratios that differ by more than 5 percent. We, furthermore, excluded the six lowest GDP per capita cities, but kept Xi'an and Chengdu for geographic completeness. This selection process yielded the following list of 15 cities for the experiment.

A-Table 3: Characteristics of Cities Used in the Online Dating Experiment

	City	GDP per capita in 2013	Urban disposable income per capita In 2013	Sex ratio of 22-32 men / 20-30 women in 2010
1	Tianjin	101689	32658	1.333
2	Beijing	92210	40321	1.210
3	Shanghai	90765	43851	1.180
4	Guangzhou	120516	42066	1.166
5	Xiamen	81572	41360	1.140
6	Shenyang	88309	29074	1.114
7	Nanjing	98171	39881	1.109
8	Hangzhou	94791	39310	1.090
9	Xi'an	57104	33100	1.078
10	Qingdao	90746	35227	1.069
11	Dalian	110600	30238	1.067
12	Jinan	75254	35648	1.037
13	Zhengzhou	68070	26615	1.031
14	Changsha	99570	33662	1.012
15	Chengdu	63476	29968	1.005

Notes: GDP per capita and disposable income data are from the National Bureau of Statistics. The local sex ratio is defined as the number of males/number of females and derived from the 2010 Census. Excluding Tianjin, the variation we have for sex ratio between the highest (1.210) and lowest (1.005) sex ratio cities for the online dating study is approximately 20 percent (0.204=(1.210-1.005)/1.005).

A-Table 4: Regression of Men's Mean Income on Local Sex Ratio with City-Level Data

Dependent variable:	Male mean income (in *log*) in a city	
	(1)	(2)
Sex ratio	0.183	0.167
	(0.498)	(0.508)
Men's income dispersion		2.041**
		(0.965)
Province dummies	Y	Y
Constant	6.970***	5.861***
	(0.036)	(0.511)
Observations	57	57
R-squared	0.582	0.673

Notes: Data are from the 2005 China mini-Census. The sample is restricted to males between the ages of 22-32 years and with an urban *hukou* and a positive income. It excluded provinces with significant minority populations and those for which we have less than 300 observations for each of men and women. The local sex ratio is defined as the *log* of the number of males/number of females. Sex ratio, mean income, income dispersion, and population size are defined at the city-level. All incomes are in *log* form. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

A-Table 5: Regression of Men's Income Dispersion on Local Sex Ratio with City Level Data

Dependent variable:	Men's income standard deviation in a city	
	(1)	(2)
Sex ratio	0.008	-0.012
	(0.132)	(0.130)
Men's mean income		0.106**
		(0.042)
Province dummies	Y	Y
Constant	0.544***	-0.197
	(0.009)	(0.293)
Observations	57	57
R-squared	0.448	0.568

Notes: Data from the 2005 China mini-Census. The sample is restricted to males and females between the ages of 22 and 32 years old with an urban *hukou* and a positive income. It excludes provinces with significant minority populations and those for which we have less than 300 observations for each of men and women. The local sex ratio is defined as the *log* of the number of males/number of females. Sex ratio, income standard deviation, mean income, and population size are defined at the city-level. All incomes are in *log* form. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

A-Table 6: Summary Statistics of Age, Income, and Education for Male Visitors

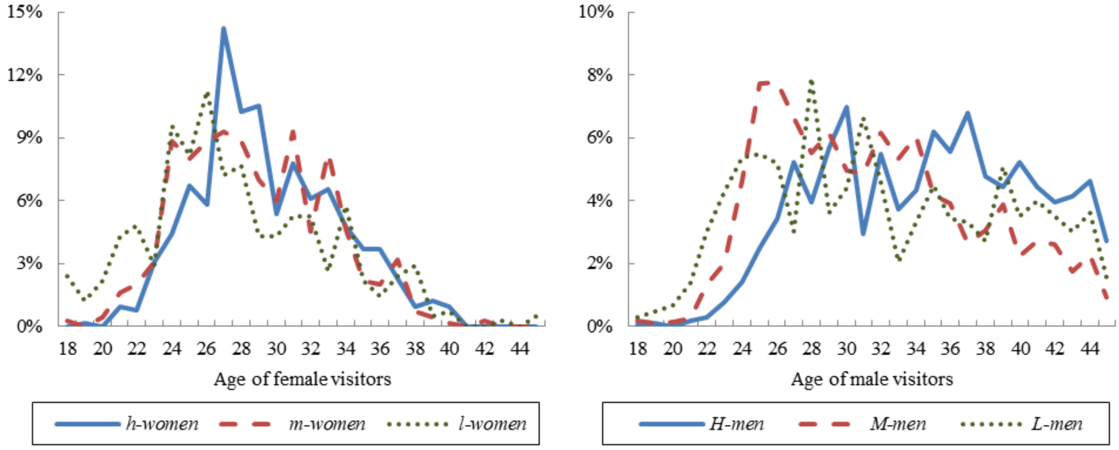
Male	Obs.	Mean	Std. Dev.	Min	Max
Age	5981	33.93	7.580	18	69
Income (1*k* CNY)	5706	10.39	11.03	1	50
Education (years)	5705	15.14	1.689	12	21

Notes: Data are based on 5,981visits from men to 390 female profiles in another experiment conducted at the same time. 275 visits did not contain income information. Among these, one did not contain education information. This leaves us 5705 visits for our analysis. Female profiles are constructed as 22, 25, 28, 31, and 34 years old, all with a height of 163 cm, a college degree, and an income of 5k-8k CNY/month. They are all unmarried with no children and are block randomly assigned to the same 15 cities.

A-Table 7: Summary Statistics of Age, Income, and Education for Female Visitors

Women	Obs.	Mean	Std. Dev.	Min	Max
Age	1811	28.86	4.405	18	45
Income (1*k* CNY)	1760	5.163	3.494	1	50
Education (years)	1760	15.54	1.387	12	21

Notes: Data are based on 1,811 visits from women to 450 male profiles in the experiment of this study. 51 visits did not contain income information. This leaves us 1760 visits for our analysis. Male profiles are constructed as 25, 28, 31, 34 and 37 years old, all with a height of 175 cm, a college degree, and an income of 3-5, 8-10, or 10-20 *k* CNY/month. They are all unmarried with no children and are block randomly assigned to the 15 cities.



A-Figure 2: Age Distribution of Women’s Visits to Male Profiles and Men’s Visits to Female Profiles

Notes: The left panel shows the distribution of women visitors to our male profiles, whereas the right panel shows the distribution of men visitors to our female profiles. We group women’s visits into three income levels: <3, 3-5, and 5-20 (in 1*k* CNY), labelled as *l*-, *m*-, and *h*-women, respectively. We group the men’s visits into three income levels: 3-5, 8-10, 10-20*k* (in 1*k* CNY) labelled as *L*-, *M*-, *and H*-men.

Appendix 3. Instrumental Variable Robustness Check

A-Table 8: The First Stage Regression for IV-Ordered Probit Regression in Table 1

Dependent variable:	Sex ratio (log)	
	For column (4)	For column (7)
	(1)	(2)
m-women dummy	-0.086*	0.028
	(0.050)	(0.029)
h-women dummy	-0.055	0.049
	(0.062)	(0.036)
Bartik sex ratio	**0.917*** **	**1.263*** **
	(0.100)	(0.063)
Bartik sex ratio**m*-women dummy	0.277**	-0.056
	(0.115)	(0.047)
Bartik sex ratio**h*-women dummy	0.214	-0.069
	(0.138)	(0.050)
Beauty ranking	-0.191**	-0.020
	(0.075)	(0.049)
Beauty ranking**m*-women dummy	0.144	-0.045
	(0.092)	(0.053)
Beauty ranking**h*-women dummy	0.069	-0.050
	(0.111)	(0.060)
Bartik sex ratio*beauty ranking	0.492***	0.008
	(0.182)	(0.105)
Bartik sex ratio*beauty ranking**m*-women dummy	-0.475**	
	(0.216)	
Bartik sex ratio*beauty ranking**h*-women dummy	-0.312	
	(0.248)	
Mean income of *H*-men		0.022***
		(0.004)
Mean income of *H*-men*beauty ranking		-0.002
		(0.007)
Mean income of *H*-men**m*-women dummy		0.006
		(0.004)
Mean income of *H*-men*beauty ranking**m*-women dummy		-0.013
		(0.008)
Mean income of *H*-men**h*-women dummy		0.006
		(0.006)
Mean income of *H*-men*beauty ranking**h*-women dummy		(0.004)
		-0.002
Additional controls:		
Age and education dummies of female visitors	Y	Y
Mean and standard deviation of men's and women's incomes in each city	Y	Y
Constant	-0.055	-0.058*
	(0.042)	(0.030)
Observations	548	548
F-statistic	115.4	167.3
R^2	0.866	0.936

Notes: The local sex ratio is defined as the number of males between the ages of 22 and 32 years old over females between the ages of 20-30 years old at the time of experiment in 2014, proxied by males between the ages of 18 and 28 years old and females between the ages of 16 and 26 years old in the 2010 Census. The *l*-women are the omitted benchmark with income less than 3*k* CNY/month. *m*-women dummy = 1 if the woman's income is between 3k and 8k CNY/month. Additional control variables are the same as those in Column (3) - (6) of Table 2. Robust standard errors clustered at the city-level are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.