# Hard to get:
# The scarcity of women and the competition for high-income men in Chinese cities[1]

20 DECEMBER 2016

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

*Reports in China of the difficulties of elite women in finding suitable mates have been increasing despite the growing scarcity of women. 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 (RDP), the pool of men the high-income women desire shrinks as their income increases, while the pool of competing poorer women expands. Moreover, for high-income women, even when high-income 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 poorer women. We test for these competitive effects with online dating field experimental, Census, and China Family Panel Studies data. Consistent with competitive entry and its deterrence, the search intensity of beautiful low-income women and high-income women—irrespective of their beauty—for high-income men increases with sex ratio and the income of high-income men. The beauty of the wife of high-income men increases with sex ratio and the men's income, as does the marriage probability of low-income women, while that of high-income women decreases. Our findings demonstrate the novel effects of women's RDP for mate income.*

*JEL* Codes: C93, J01, J12

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

---

[1] We thank numerous scholars and participants at the Tokyo Labor Economics Conference 2016, the NBER Conference China Economy Working Group Conference 2016, CUHK Economics Department Workshop on Family and Labor Economics 2016, Asian Meeting of the Econometrics Society 2016, and Chinese Economic Association Meeting 2016 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 (Y. Wei, Jiang, and Basten 2013; You, Yi, and Chen 2016) of the difficulties of elite women in finding suitable mates have been increasing, despite the growing scarcity of marriageable women in China (Jiang, Feldman, and 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 (Zhu, Lu, and Hesketh 2009).

One might have supposed that the surviving women can only benefit from their own 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). Already, evidence exists for the expected increase in competition among men or their supporting families and the relaxation of competition among women. The rise in local sex ratio (population of men/population of women within a province or city) predicts not only increases in men's level of entrepreneurship (Yuan, Rong, and Xu 2012), men's work hours in dangerous and risky jobs (S.-J. Wei and Zhang 2011), male criminal activities (Edlund et al. 2013), the savings of families with sons (S.-J. Wei and Zhang 2011), the time men spend on housework within households (Du, Wang, and Zhang 2015), and women's participation in decision making (Edlund et al. 2013); but also decreases in women's educational attainment, and employment (Edlund et al. 2013). These studies about outcomes, which primarily focus on the behavior of men, confirm Becker's (1973) theory that the bargaining position of women improves as sex ratio rises.

Moreover, recent evidence indicates a preference basis for these findings of an increasing competitiveness gap between men and women in the context of high sex ratio. Ong and Wang (2015) reveal an asymmetry in the preference for mate income in a field experiment on a large online dating website. They use random assignment of incomes to artificial profiles to rule out factors other than income [e.g., grooming style, height (Ong 2014), "chemistry," or meeting opportunities (Nielsen and Svarer 2009)] as the cause of revealed preferences. They report that although men of all income levels are largely indifferent to women's income, women of all income levels prefer high-income men to low-income men.[2] Such an asymmetry in preferences for mate income gives men (and their families) an extra mate attraction incentive to achieve higher income levels (or wealth) as women become more scarce.

However, and surprisingly, Ong and Wang (2015) find that women's preference for mate income increases with their own income; their preference for mate income is reference-dependent. This unexpected finding is, nonetheless, consistent with a relative income preference reported in an empirical study of American online dating data (Hitsch, Hortaçsu, and Ariely 2010). Moreover, this reference-dependence takes a straightforward interpretation. A woman would naturally seek out a husband whose income would more than offset the opportunity cost of her (potentially) lost income from decreased labor market participation after marriage or childbirth (Lundberg and Rose 2000; Waldfogel 1997) particularly if she "opts-out" completely (Hersch 2013). This notion that a woman might choose a mate with a view to offsetting her opportunity cost is consistent with both the long standing theory of habit formation and with the recent behavioral theory of reference-dependent preference in which the reference point is lagged consumption (Kőszegi and Rabin 2012) or the difference between the income of the woman's husband and that of her sister's husband (Neumark and Postlewaite 1998). Her candidate husband may also anticipate her decreased labor market participation after marriage and discount her expected contribution to household income accordingly. Therefore, traditional gender roles can induce reference-dependent preference on the part of women, as well as an asymmetry in preferences for mate income between men and women.[3] As we will make evident, such a "preference" (which we will adopt as a primitive notion), induces heterogeneity among women and a counterintuitive comparative statics, one consequence of which is that high-income women can be worse off when women in general are scarcer.

To begin, while the scarcity of women might predict in a contest framework that men would work harder (Clark and Riis 1998a, 1998b) or take greater risks to be attractive to women, the subsequent increase in their "mate prize value" in the context of a

[2] In the experiment, the education of the online dating profiles was held constant at the college level.

[3] Although we only included employed wives in our Census data, we do not know if these wives had reduced or planned to reduce their labor market participation, and thus, cannot rule out this possibility.

homogenous preference for mate income on the part of women would predict that *all* women will increase or decrease their search intensities for these even higher high-income men: increase because the returns to effort are higher, and decrease because effort is less necessary. Similarly, if men become relatively more plentiful than women, then the increase in the number of prizes for women predicts that *all* women enjoy more opportunities to date higher income men. This effect can, also, either increase or decrease search intensity for *all* women.

On the other hand, a reference-dependent preference (RDP) for mate income on the part of women can induce heterogeneous behavior; such preferences should engender a *nested* prize structure. High-income women will compete with all women for high-income men. By contrast, low-income women would compete only with each other for low-income men. For illustration (in the extreme case), the prize of high-income women is the prize of both high- and low-income women, but the prize of low-income women is not a prize for the high-income women. More generally, women's RDP for mate income implies that high-income women are less willing than low-income women to settle for low-income men when the competition for high-income men increases. The key characteristic of this RDP is to intensify the competition women face as their income increases by reducing the pool of men they desire while expanding the pool of other women who desire these men. The fierceness of the competition for men with even higher income than herself "escalates" as a woman's income increases.

This escalation in competition can, moreover, be exacerbated by the increases in the income and the availability of high-income men because either may boost the expected return of pursuing such men: the former increases the value and the latter increases the probability. Thus, the *direct effects* of both 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 sex ratio, there is a greater number of high-income men for each woman to desire. But, secondly, in either case, the higher expected returns for pursuing these high-income men may also increase the number of low-income women 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. Thus, 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. That makes high-income women, who are averse to matching with low-income men, worse off. The indirect effect is more likely to dominate the direct effect for high-income women, while the opposite is true for low-income women, who enjoy more options, given that they can be satisfied with matching with low-income men.

In summary, the interaction of increases in high-income men's desirability due to increases in their incomes and their availability with women's RDP for mate income, *further escalates* the competition that high-income women face. Consequently, high-income women can be made worse off when high-income men are even richer or more plentiful, because both increase the returns to poorer women in pursuing these men. Such may be the situation in China, where sex ratio and men's income compared to women's have both been increasing dramatically over the last 30 years (Gustafsson and Li 2000).

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 (Becker 1973). We exploit variation in local sex ratio and the incomes of men across Chinese cities to test for these comparative statics effects of escalating competition on the online dating search intensity and marriage probability of women with different income levels. We employed an online dating field experiment across 15 major cities to measure variations across women's search intensities for men. In this experiment, we randomly assigned three income levels to 450 artificial male profiles on a large online dating website (with more than 60 million members in 2011) and recorded the incomes and other characteristics of 1,811 "visits" from women to these male profiles, with the women being divided into high-, medium-, and low-income levels. We also had nearly two-thirds of the profiles of these female visitors rated for their beauty.

We confirm the RDP finding that women's search intensities for high-income men, as measured by the probability of visits to these male profiles, increase with the women's own incomes. Importantly for this study on the comparative statics of women's RDP, we show that the search intensity of the beautiful low-income women for high-income men increases with sex ratio and the income of high-income men, while the search intensity of the plain among these women decreases. In reaction to the competitive entry of these low-income women, the search intensity of high-income women--irrespective of the beauty--for the high-income men increases on both local sex ratio and the income of these high-income men. This stronger reaction among high-income women is expected if women's RDP for mate income makes them averse to settle for low-income men when the competition for high-income men increases. Although an increase in the income of high-income men can also increase the competition for them even if women have absolute/non-RDP for mate income, such a non-RDP does not predict the strength of the women's response increasing on the women's own income.

The consequence of high-income men enjoying a larger pool of more attractive women to choose from is evident in the China Family Panel Studies (CFPS) data set. The beauty of the wife of high-income men increases on sex ratio and their own income. Although online dating data show that the search intensity of these high-income men for beautiful women increases with their own income, it does not increase with sex ratio. Thus, our evidence suggests that this increase in the beauty of their wife when sex ratio increases is not due to an increase in the effort of high-income men to acquire a beautiful girlfriend/wife when sex ratio rises, i.e., when they face more competition from other men.

The expected consequence of the increased entry of low-income women into the market for high-income men is evident in Chinese Census data. While we confirm the standard result that the marriage probability of women decreases with their educational attainment (Boulier and Rosenzweig 1984), we also find the novel result predicted by women's RDP that only the marriage probability of high-income women decreases significantly with local sex ratio and the income of high-income men, notwithstanding the increase in these women's search intensity. By contrast, low-income women's marriage probability increases significantly on sex ratio and weakly on the income of high-income men, despite their aggregate (irrespective of their beauty) search intensity for these men not significantly increasing on either. The contrast in marital outcomes between high- and low-income women in either case of increases in sex ratio or the income of high-income men demonstrates the dominance of the effect of a greater number of women desiring the same men over that of a greater number of desirable men per high-income woman. The reverse is true for low-income woman, as predicted by women's RDP. While the greater search intensity of high-income women for high-income men, when either sex ratio or the income of high-income men increases, might be predicted by positive assortative matching, the decrease in these women's marriage rate, both relative to low-income women and in absolute value, is not.

Our results based on the income levels of women can be biased by their decision to participate in the labor market. We find similar qualitative results when we impute the wages for women using their age, educational attainment, and the number and gender of children according to the methodology in Zhang and Liu (2003). (These results are available on request.)

Our findings demonstrate the novel effects of women's RDP for mate income and could help explain the increasing difficulties of high-income women in finding mates in China, notwithstanding the increasing scarcity of women. Despite the novelties 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 on sex ratio.

Though our findings support prior studies of high-income women's RDP for mate income, our main contribution is the use of variations in local sex ratio and men's income to test for the novel comparative statics implications of women's RDP. The heterogeneity in women's behavior and outcome, according to their beauty, level of income, and the interaction of the two, is the main basis by which we identify RDP. To our knowledge, such heterogeneity in behavior and outcomes (even to the extent that some women benefited from while others are hurt by increases in sex ratio and men's income) is not predicted by established theories, nor is it, as far as we can tell, consistent with previously tested confounders.[4] Although several studies have examined loss aversion in contests (H. Chen, Ham, and Lim 2011; Z. C. Chen, Ong, and Sheremeta 2015; Ernst and Thöni 2013), we are the first to model and test for the comparative statics of RDPs in real life contests outside the laboratory.

## II. Hypotheses

One method for modeling the reference-dependent *component* of the wife's utility is

[4] Laboratory experimental and empirical finance studies on RDPs based on the pioneering work of Tversky and Kahneman (1991) abound. We contribute to the small but growing empirical literature, outside of finance, that aims to identify RDP (Barberis 2013). In addition to showing direct evidence of RDP in dating and marriage data, we are the first to demonstrate the comparative statics effects of such preferences.

to simply incorporate the difference in her income and that of her husband into her utility:

$$U_{wife}\left(I_{husb}-I_{wife},\dots\right)=U_{wife}(other\ goods)+f(I_{husb}-I_{wife})+\dots.$$

where $f$ increases on the difference, such that it is positive when the difference is positive and negative otherwise.[5] Our conceptual framework focuses on how this reference-dependent component influences the competition between the high- and low-income women for high-income men.

Figure 1 illustrates how the competition that a woman faces for men whose income is higher than her own "escalates" with her income, due to women's RDP for mate income. We show how this escalating competition interacts with increases in the desirability (as indicated by income) and the availability (as indicated by sex ratio) of all men including high-income men to decrease high-income women's odds of matching.

The horizontal axis of Figure 1 represents a ranking of individuals of each gender by income from lowest (LHS) to highest (RHS). The vertical axis represents the relative frequency of individuals with a specific level of income.

[Insert Figure 1 here.]

To define our key concept of escalating competition, consider a woman who earns $f_0$. Given RDP, she prefers men who earn more than she does ($f_0 < m$) over those who earn less. For any of these men who earns $m_0 > f_0$, the woman who earns $f_0$ is competing against all of the women who earn less than him ($f < m_0$), due to other women's RDP.[6] The crucial implication of women's RDP is that the fierceness of the competition this woman faces increases with her income ($f_0$), which shrinks the pool of men who earn more than her ($m > f_0$), while it expands the set of women who earn

---

[5] $f$ can be expressed as follows: $f\left(I_{husb}-I_{wife}\right)=\begin{cases} f\left(I_{husb}-I_{wife}\right), if\ I_{husb}\geq I_{wife} \\ -\lambda f\left(I_{wife}-I_{husb}\right), if\ I_{husb}< I_{wife}\end{cases}$, where $\lambda$ indicates the degree of the aversion to the wife having a higher income than the husband. $f$ does not have to be discontinuous. Farber (2008) uses a function that with a discontinuous marginal utility at the reference point. Such RDP reflected in the utility function can emerge from underlying preferences or other sources. In the latter case, the utility function is an indirect utility function. Later, we will argue that RDP for mate income on the part of women can arise from their attempt to cover the anticipated opportunity cost of lower labor market participation after marriage, following traditional gender norms. The specific functional form or the size of $\lambda$ is not important for our purpose, which is the empirical identification of the comparative statics effects.

[6] Note that though her income determines her reference point, and her reference point determines whom she is more likely to accept, her reference point does not completely determine the degree of her competition. That is determined by the income of the man she is competing for.

less than any of these men ($f < m_0$). Thus, the competition[7] that a woman faces for higher income men "escalates" on her income because of women's RDP.

We now explain how the increase in men's income or sex ratio can interact with the competition that escalates with a woman's income to further escalate the competition that she faces. When men's income increase, either the solid (blue) distribution shifts to the right, or more mass is distributed to the right. In either case, the set of higher income men ($m > f_0$) expands, which benefits the woman who earns $f_0$, because it increases the expected return of pursuing these men. However, a second effect is that it increases the expected return of all women who earn less than $f_0$ in pursuing these men. This latter effect increases the share of all women who might pursue or accept an offer from each of these men (e.g., from 40% of the women who earn less than $f_0$ to 60%) by making these richer men preferable to poorer men with better non-income qualities, and thus, increasing the competition for these richer men.

The marginal impact of the increase in the desirability of high-income men is likely larger for high-income women than for low-income women, who consider low-income men acceptable options. Hence, we expect a greater increase in the search intensity/attraction effort of high-income women when the income of high-income men increases than that of low-income women. Nevertheless, the small increase in effort (including "entry") of a large population of lower income women may overwhelm the large increase in effort of a small population of high-income women, and consequently, crowd them out of the mating market.

This effect of the increased effort of a larger population of lower income women in the competition for high-income men can be magnified if women are heterogeneous in a characteristic that men care about, e.g., beauty. The large population of lower income women is likely to have a greater number of outstanding beauties. When high-income men's income increases and become more desirable, these women who may initially have preferred more (e.g., physically) attractive but poorer men, may switch to preferring less (e.g., physically) attractive but richer men, even if that entails more effort

[7] We assume that women can compete through search intensity or level of effort in being attractive to these high-income men. However, we also include passive choices that can have a competitive effect, e.g., rejecting these men less. The latter can be the case, if the increased mate value of these men with respect to the income dimension can more than compensate for deficiencies in other dimension (e.g., height and looks). In either case of active or passive competition on the part of women, fewer high-income men remain in the market for other women to find or attract.

by these women

Sex ratio can exert a similar negative effect on the mating prospects of high-income women with high-income men. Though an increase in the availability of high-income men does not increase their desirability, it can increase the expected return of pursuing these men for low-income women, because it increases the probability of winning one of these men. The overall sex ratio, which can increase because there are more low-income men, need not take this possibility of increased entry into account. Even if the overall sex ratio is in a woman's favor, the income-specific sex ratio, which can be defined as the number of men who earn more than her ($m > f_0$) over the number of women who earn less than her ($f < f_0$), may not be in her favor because of women's RDP, if she earns a high enough income. What is required for sex ratio to predict the direction of effect of the availability of high-income men on the mating prospects of high-income women is that the share of high-income men increases with sex ratio, which will be shown to be the case in our data..

Although our analysis up to this point focuses on individual men and women, it can apply equally to men and women of different income groups: top-1/3 (high), middle-1/3 (medium), and bottom-1/3 (low) of their respective populations. Again, while the overall sex ratio may favor the high-income women group, the income-specific sex ratio may not. Analogously with individuals, in the case of groups, we can observe the potentially detrimental effect of the increase in the income, and therefore, desirability of high-income men if the mean income of the high-income men group is higher than the mean income of the high-income women group to begin with. Given this income gap, we can use changes in sex ratio to observe the potentially detrimental effect of the increase in the availability of high-income men, if as with the analysis of individuals, the availability of high-income men increases with sex ratio.[8]

In fact, we show with online dating data and with Chinese Census data that the income of high-income men in China is higher than that of women of all income levels, including high-income women. Moreover, our data show not only that men's incomes are positively correlated with sex ratio, but also that the correlation is due to an increase

[8] The income specific-sex ratio (the number of high-income men over the number of all women) is proportional to (a third of) the sex ratio for income groups based on fixed shares of a population. Thus, in regressions with income groups, the coefficient for sex ratio would be some proportion of the coefficient for income specific sex ratio.

in the share of high-income men. Therefore, sex ratio increases with the proportion of men with high income to women of all income levels on both the online dating website and in the surrounding city. Thus, in the case of China, sex ratio alone can be used as a treatment variable to test for the level of competition that women of all income levels might face for high-income men both in their search on the online dating website and in their offline probability of marriage. We can use the interaction between sex ratio and the high-income women dummy variable level to test for the comparative statics effects of the changes in the availability of high-income men on the competition for them that we posit to escalate with the women's income.

When there are more high-income men or when the income of the high-income men increases, there will be two effects on women of all income groups. The direct effect is that there will be a greater number of desirable men, when sex ratio increases, or men are more desirable, when their income increases. In either case, the expected value of pursuing these men increases for all women. A consequent indirect effect in either case is that there could be a greater number of women desiring the same men. Our prediction for high-income women is that the direct effect will likely dominate the indirect effect. Our prediction for low-income women is the opposite. High-income women are likely to be worse off in terms of marital outcomes when sex ratio or the income of high-income men increases. In contrast, low-income women are likely to be better off. We focus on the highest and lowest income categories of women because the two effects are offsetting, and thus, indeterminate in their net effect on the medium-income women.

We further illustrate the crucial aspects of these intuitions with a game theoretic model in Appendix 1. It shows in a numerical two-player example how the increased desirability or availability of high-income men can hurt high-income women through the potential increase in the efforts of low-income women to match with these men. In particular, A-Figure 1 shows the probability of the low-income woman choosing $Try$, which we here interpret as the share of the population[9] of low-income women choosing $Try$, jumping at $\theta = 1 + c$, when the value of high-income men ($\theta$) is high enough for

[9] The probability with which individual players choose an action in a mixed strategy 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

the low-income women to relinquish their low-income men option. However, this share of low-income women choosing $Try$ decreases steadily and quickly crosses from above to below the share of the high-income women choosing $Effort$. The difference between the share of high-income women choosing $Effort$ and the share of low-income women choosing $Try$ increases as the value of the high-income men increases up to $\theta = \frac{1+c}{z}$, which is also the point at which the share of low-income women choosing $Try$ approaches its minimum within the (realistic) region in which both types of women are not unanimously choosing one strategy (i.e., choosing mixed strategies with non-degenerate probabilities). Thus, we expect that the increase in the search intensity of high-income women will become more significant as the expected value of pursuing high-income men grows large, especially as they have no recourse to low-income men, whereas that of low-income women will become less statistically significant, even to insignificance, all the more so due to the following inherent limitation of our design.

We created a fixed number of high-, medium-, and low- income profiles across all cities of the experiment. Thus, in cities with more high-income men, our high-income profiles are likely to receive a smaller share of all visits to high-income male profiles. Therefore, our test for increases in the share of women's visits to our high-income male profiles in cities where high-income men have higher incomes (including higher sex ratio cities, since these cities have a greater share of rich men) will be biased towards false negative coefficient findings or merely false insignificant findings. Since this possibility of "false negative" findings is most likely to affect our measurement of the change in the lower search intensity of the low-income women, we formulate our first hypothesis in terms of high-income women's search intensity.

***Hypothesis 1.*** The search intensity of high-income women for high-income men can increase significantly relative to low-income women (or even absolutely) as sex ratio or the income of these men increases.

Although the increase in the search intensity of low-income women may not be sufficient to overcome the bias towards false negative findings in our experiment, the increase in desirability and availability of high-income men can, nonetheless, induce a measurable increase in the search intensity of the beautiful among low-income women. Therefore, we predict that

***Hypothesis 2.*** The search intensity of beautiful low-income women for high-income men can increase significantly relative to plain low-income women (or even absolutely) as sex ratio or the income of these men increases.

We group our hypotheses for conceptual clarity in terms of the type of effect, but we group our corresponding observations for efficient exposition in terms of the treatment variables. We first report the results related to increases in sex ratio in Observation 1 and then report those related to increases in the income of high-income men in Observation 2.

As a consequence of the entry of attractive low-income women into the competition for high-income men,

***Hypothesis 3.*** The beauty of the wife of high-income men can increase with sex ratio or the income of the men.

Given the entry of low-income women (regardless of beauty) into the competition for high-income men,

***Hypothesis 4.*** The marriage probability of high-income women can decrease relative to low-income women (or even absolutely) with the increase in sex ratio or the income of high-income men.

Thus, we predict what may otherwise appear to be a set of paradoxical outcomes; that high-income women can be hurt when the high-income men they seek become richer and paradoxically, more available, while high-income men benefit when they face more competitors among themselves and men of other income levels. However, even this latter result is expected if high-income men must become more popular to *all* women, and the consequent increase in demand for these men dominates the effect of an increase in their supply.

## III. Online Dating Field Experiment

### *A. Experimental Design*

Our field experiment extends the work of 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 labor market discrimination. We used one of the largest online dating websites in China, with a reported membership of 60 million members in 2011. 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" [10] of their profile. 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, income, and the first two lines of a free-text statement. Users can then click a link and "visit"[11] 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 about 6 CNY. We only recorded 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.[12] These profiles were posted for only 24 hours, after which, the accounts were closed. To further minimize any

[10] 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.

[11] Visits are a credible measure of preferences. Although visits are free, they involve the opportunity cost of time that can be spent on visits to other profiles. Their relative frequency over male profiles with randomly assigned incomes should reveal preferences for those incomes. Visits without a follow-up email also do not imply an offer that can be rejected. They can only be based on abbreviated profiles that contain one picture, income, age, city, and two lines of the personal statement. At best, the decision to click and visit expresses an *ex-ante* interest, which may not be sustained *ex-post* after viewing more of the profile. Thus, we do not expect visits to be made strategically to avoid humiliating rejections.

This study focuses on the effect of the scarcity of women on their search intensity given the RDPs identified in Ong and Wang (2015). We refer the interested reader to their paper for a full discussion on how visits measure preferences for mate income.

[12] We are unaware of legal restrictions on the noncommercial 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 like their European counterpart, 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 here: http://www.virginia.edu/vpr/irb/sbs/resources_regulations_subparta.46.101.html#46.102(i) 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[13], which we will call low-, middle-, and high-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,[14] 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[15] and the marital status of "single with no children" and "buy a house after marriage" (i.e., did not already 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[16] 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 only records visits to individual profiles once from any visitor. Visits across profiles are not necessarily from unique visitors. However, random assignment of characteristics to our profiles should rule out the individual idiosyncratic factors of our visitors as the main driver of our findings.[17]

The website offers a number of methods for ranking the profiles of other users in its search engine, including registration time, login time, age, number of photos, credibility of the profile, and income. Given the random or constant assignment of all

---

[13] These income levels were calibrated based on pilot experimental data. We chose these levels of men's incomes to be high enough to receive a sufficient number of visits from women within a short period of time without being conspicuously high. These levels are similar to those chosen in Ong and Wang (2015). Supporting the rationale for our choice, the income level that we chose for low-income men (3-5*k*/month) was slightly lower than what female respondents said was satisfactory (6*k*/month) for a mate in a national survey three years later:
http://www.scmp.com/news/china/society/article/1913694/great-expectations-chinese-womens-ideal-man-should-earn-6701-yuan.

[14] The website's software focuses the profile picture on the face. Therefore, such pictures should not exhibit clues about income, height, or characteristics that could conflict with those we assigned them.

[15] According to CFPS data for 2010, the average income for male 30-40 years old college graduates in Beijing and Shanghai is around 8k/month. The lowest income profiles in our experiment is 3-5k/month, which is roughly between the 20th and the 50th percentiles. However, the sample is quite small (44 observations for male 30-40 years old college graduates in Beijing and Shanghai). We do not have comparable Census data for income distribution in 2010.

[16] We were prepared to carefully eliminate any possible inconsistencies between statements and other parts of the reconstructed profiles, although we did not find any.

[17] 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.

characteristics, these different ranking methods should not affect on our results.

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 5 minutes between any 2, extending to at least 10 minutes between any 2 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 total number of visits our profiles received.

In total, our male profiles received 1,811 visits from females. Among them, 1,474 have photos. We subsequently rated a random sample of two-thirds of these female visitor's photos for their beauty using a proprietary rating program which raters can access 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. Then each $i$ photo is paired a second time to be rematched itself after being put back into the pool. Thus, each photo is matched a total of 20 times. Each photo was on average rated 200 times, which is approximately ten times the frequency of other studies (Deryugina and Shurchkov 2015). In total, 692 Chinese raters (326 male) rated 867 photos. The raters were graduate students from Peking University HSBC Business School recruited through a mass email. We used two rounds (one-third of photos in each round), because of our limited capacity to recruit raters in the first round. We paid raters 5 RMB[18] to rate100 pairs of photos in the first round (January 4, 2016) and 10 RMB to rate100 pairs of photos in the second round (July 3, 2016).

We asked raters to choose the more physically attractive within each pair of 100 pairs instead of asking for a numerical rating within a certain range of numbers, as is standard in the field (Hamermesh and Biddle 1994). This binary judgement 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

[18] At the time of writing, the exchange rate was 1 USD for 6.5 RMB. Given the few minutes it takes 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.

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 (Ong, Yang, and Zhang 2016). 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. We utilize the reported incomes of the male visitors attracted by these female profiles to construct the distribution of men's income in the 15 cities of our main experiment on the website.

### *B. Overview of the Raw Experimental Data*

Before we present the regression findings, we reveal suggestive evidence in A-Figure 2 of Appendix 3 from our raw data. The figure shows that the means of the men's incomes increase with sex ratio for the 15 cities that we used for the experiment on both the website and in the surrounding city population. This result is important to establish our interpretation of our main findings, which is that if the share of visits received by our high-income male profiles, compared with our low-income male profiles, increased with sex ratio, then we can infer that our high-income male profiles were visited with higher probability than our low-income male profiles. We show first that this inference follows, in particular, when we have the same number of profiles (10 for each of our three income levels) across all cities.

The left side of A-Figure 2 is based on the reported incomes of 5,535 visits from men aged 18 to 45 to the 390 female profiles in the experiment we conducted simultaneously in the same 15 cities (Ong, Yang, and Zhang 2016). Based on city-level population data from the 2010 China Census, we grouped these cities by local sex ratio into top-, medium-, and bottom-five-city groups. The graph shows that the distribution of high sex ratio cities (top-five-city group) is more right-skewed than that of those in the medium- and bottom-five-city group. The right side of A-Figure 2 displays a similar pattern for the top-third of the 243 cities in the China 2005 1 percent Population Survey. This finding is supported by regression results in A-Table 4 and A-Table 5 in Appendix

3. Thus, the income of men both on the website where we conducted the experiment and in the surrounding city population increases with local sex ratio. This implies that increases in sex ratio are not driven by a disproportionate increase in the share of low-income men with respect to high-income men. Therefore, we can conclude that *there are more rich men and/or men are richer both on the dating website and in the surrounding city in higher sex ratio cities*.

Recall that we fixed the number of profiles (10 for each of our three income levels) across all cities in our online dating experiment, which is only a small part of this large online dating website. Thus, our high-income profiles should have received a smaller 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 below. *Our high-income profiles can only receive a larger share of all of 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.*

This increase in women's visits is already evident in the graphs of our raw data in Figure 2. The graphs exhibit visits by women[19] to male profiles in three groups of cities: top-, middle- and bottom-five, according to the local sex ratio of the 20 to 29 age cohort in the 2010 Census (24 to 33 age cohort at the time of the experiment), from the highest local sex ratio to the lowest. The horizontal axis indicates the ages of our male profiles. 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 5-city group divided by the visits to all our profiles over all male incomes types in the same five-city group. The pattern suggests that the marginal impact of increasing the incomes of men from middle-(8-10*k*) to high-(10-20*k*) incomes on the visits of women increases on local sex ratio.

[Insert Figure 2 here.]

The summary statistics of age, income, and education for each gender of our visitors are in A-Table 6 and A-Table 7 in Appendix 4.

---

19 This website has no option for users to report a same sex preference, but users can view anyone else's profile. We did not receive any same sex visitors.

We grouped female visitors into three income levels: <3, 3-8, and 8-20 (in 1*k* CNY). These are labeled as *l*-, *m*-, and *h-women*, and are represented by three lines, respectively in Figure 3. 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 earn less than men. Both the left (low sex ratio cities) and the right (high sex ratio cities) panels of Figure 3 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 behavior, but rather on the effect of men's income relative to women's incomes on women's behavior 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 own reported incomes.

[Insert Figure 3 here.]

Second, we show a kink in their graph of the high-income women at point B and B' suggesting 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*), i.e., as the profile's income exceeds the women's average income (which is roughly 15*k*). Third, previewing the main findings in Table 1, Figure 3 shows 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 that 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'). *h*-women also made roughly 75 percent of their visits to the 10-20*k* male profiles in bottom-8 sex ratio cities (C), but approximately 85 percent of their visits to that type of profile in the top-7 cities (C'). Together, these three levels of evidence already suggest the increased search effort of high-income women, due to women's RDP, even before we impose controls econometrically.

### *C. Regression Analysis*

In this section, we confirm the impression from our raw data by formally testing Hypothesis 1 which predicts that the search intensity of high-income women for high-income men should increase at a relatively higher rate than that of other types of women as sex ratio or high-income men's income increases.

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 15 cities. 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 those are made to one of our three income types of male profiles. We treat each as one of the three choices in an ordered logit 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.[20]

We group visits from women into three income levels: <3, 3-8, and >8 (in 1*k* CNY), and associate a dummy variable with each level: *l*-, *m*-, and *h-women*, respectively. The low-income level is the omitted benchmark.[21] We calculate the local sex ratio using county-level data based on the full sample of the 2010 Census.[22] The 2010 Census released only the aggregate number of people of each gender in five-year age groups. Thus, we use the sex ratio for individuals aged 20-29 (who were 24-33 years old at the time of the experiment). As a robustness check, we also use the sex ratio of those aged 25-34 and 20-34 (who were 29-38 and 24-38 years old, respectively, at the time of the experiment) and find similar results.

The individual characteristics of our women visitors that we collected include income,

---

[20] Note that our treatment variable in the experiment is men's income type (*H, M, 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.

[21] We used absolute cutoffs for incomes in the online dating section of the study because the website aggregates incomes into nine levels.

[22] See the tabulation of the 2010 Population Census at the County Level by the National Bureau of Statistics

age, years of education, and height. We calculate the wage distribution (means and standard deviations of both men's and women's income) from the incomes of our visitors. We collect city characteristics, namely, GDP per capita, and migration share (the share of population without local *hukou*[23] of the total population in a city) from the 2010 Census data.

The ordered logit regression models the probability that a woman from a specific level of income visits a male profile of a specific level of income among all income levels of male profiles. We interpret this probability as 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. We control for city-level income, 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.

Table 1 exhibits the results of the ordered logit regression of women's visits as a function of their own income and local sex ratio. The positive terms for the *h-women dummy* (1.780) in column (1) indicates that high-income women visit high-income male profiles more than low-income women, confirming our impression from Figure 3 and supporting previous findings (Ong and Wang 2015). Column 2 of Table 1 shows not only that the intercept for high-income women is higher (1.324) than that of low-income women, but also that the difference increases on sex ratio (8.829).

[Insert Table 1 here.]

Importantly for our competitive entry hypothesis, the coefficient for *sex ratio* for the benchmark low-income women is small and statistically insignificant in columns (2)-(4) of Table 1. This insignificance can be due to our design being naturally biased toward 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 both richer and more plentiful. Hence, the reader should perhaps interpret the weakly negative coefficients as weakly positive. However, the negativity of the coefficient or the lack of significance can also

[23] An internal passport system from the command economy era (Chan and Buckingham 2008): a *hukou* entitles holders to government social services, which are often paid for by employers

be due to low-income women enjoying more outside options among low- and medium-income men, respectively. Indeed, the probability of $Try$ for low-income women is lower than the probability of $Effort$ for most of the region of the mixed strategy equilibrium in A-Figure 1 of the theoretical example in Appendix 1. Nonetheless, the negative coefficient or lack of significance can in addition be 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 those beautiful low-income women for these same men.

In contrast to low-income women, columns (2)-(4) show that the coefficient for high-income women interacted with sex ratio is significantly positive. This finding indicates that the search intensity of high-income women for high-income male profiles increases relative to that of other women and even absolutely when local sex ratio is high. We calculate the marginal effects of sex ratio on high-income women's probability of visits based on the coefficients of the ordered logit regression in column (5) of Table 1, keeping all variables at their mean values. The 28.759 coefficient of *sex ratio*h-women dummy* indicates that a 10 percent increase in the sex ratio increases the probability of high-income women visiting high-income male profiles by 10.09 percentage points, and decreases the probability of visiting middle- and low-income male profiles by 6.60 and 3.49 percentage points, respectively, compared with that of *l*-women.[24] These marginal effects are also very close to those corresponding to column (2) of Table 2.

These findings are highly consistent with our theoretical results as shown in the top part of A-Figure 1 in Appendix 1. First, our online dating results should be modeled by the top part of A-Figure 1, after the crossing of the effort levels of high- and low-income women, because the high-income men (10-20*k*) earn considerably more

---

[24] In our ordered logit model, the probability of each type of male profile being visited is given by $P(L=1)=\frac{1}{1+exp(X\beta-\kappa_1)}$, $P(M=1)=\frac{1}{1+\exp(X\beta-\kappa_2)}-\frac{1}{1+exp(X\beta-\kappa_1)}$, and $P(H=1)=1-\frac{1}{1+exp(X\beta-\kappa_2)}$, where $\kappa_1$ and $\kappa_2$ are the estimated cutoffs. 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 $\beta_i$ of $X_i$, the marginal effect $\frac{\partial P(L=1)}{\partial X_i}=-\frac{\beta_i exp(X\beta-\kappa_1)}{[1+exp(X\beta-\kappa_1)]^2}<0$, $\frac{\partial P(H=1)}{\partial X_i}=\frac{\beta_i exp(X\beta-\kappa_2)}{[1+exp(X\beta-\kappa_2)]^2}>0$, wheres $\frac{\partial P(M=1)}{\partial X_i}=\frac{\beta_i exp(X\beta-\kappa_1)}{[1+exp(X\beta-\kappa_1)]^2}-\frac{\beta_i exp(X\beta-\kappa_2)}{[1+exp(X\beta-\kappa_2)]^2}$ is in general ambiguous.

than the low-income men (3-5$k$). Second, A-Figure 1 shows, as we find empirically, that when high-income women's effort is high (i.e., high probability of $Effort$) and increasing on the value of high-income men, low-income women's effort should be low (i.e., low probability of $Try$) and decreasing.

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 estrogen, and therefore fertility (Rhodes 2006). However, with few exceptions, facial beauty is generally 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 (Ong, Yang, and Zhang 2016). In this article, we merely note that beautiful low-income women may still expect good odds of matching with high-income men even when the competition for these men increases (see Appendix 1 for details). Therefore, we control for facial beauty in column (5) of Table 1. Notably, though we control for women's facial beauty, which is highly correlated with their income (0.295, p=0.005 in a simple regression of *log* women's income on beauty), the interaction of *sex ratio* and *high-income women dummy* remains significant (28.759).

Importantly for our previous finding of a lack of significance in the increase in the search intensity of low-income women, column (5) of Table 1 also reveals heterogeneity in the reactions of women with different income levels to increases in the sex ratio, according to their beauty. The highly significant negative coefficient for *sex ratio* (-16.657) indicates a pronounced decrease in the search intensity among the plain of the benchmark low-income women for high-income men when the sex ratio increases. By contrast, the highly significant positive coefficient for *sex ratio*beauty* (30.568) indicates a pronounced increase in the search intensity among the beautiful among low-income women for high-income men when sex ratio increases. Thus, increases in sex ratio induce divergent reactions among beautiful and plain 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), although the high-income women appear to be defending against greater competition.

The medium- and high-income women show diminishing contrasting reactions by their beauty as sex ratio increases, due to their RDP. The significant positive coefficient (12.624) for the interaction between sex ratio and the medium-income women dummy suggests that the plain among them react less negatively (-16.657+12.624) than low-income women (-16.657) to the increase in sex ratio. In contrast, the significant negative coefficient (-21.539) for the interaction among sex ratio, beauty, 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 (30.568-21.539) than those among low-income women (30.568). The positive and significant coefficient (28.759) for the interaction of sex ratio and the high-income women dummy suggests that the plain among the high-income women search more intensively for high-income men when these men are more plentiful. Similarly, in contrast, the negative and insignificant coefficient (-48.232) for the high-income women suggests that their reaction to the sex ratio is insignificantly influenced by their beauty compared to that for low-income women. Again, this pattern of decreasing differentiation in search intensity between women of different income levels by their beauty, 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.

To summarize, the reaction of the plain among medium- and high-income women is strictly less negative than that of the plain 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 behaviors of the high-and low-income women when sex ratio increases is between the plain high-income women (28.759) and the plain low-income women (-16.657). A lack of beauty makes less difference to the search intensity for the plain high-income women when sex ratio increase than for the plain low-income women. This result is expected if high-income women are more desperate (less willing to avail themselves of the option of low-income men) to match with a high-income man, when the competition for them increases, due to their RDP.

***Observation 1.*** The visits of high-income women and the beautiful low-income women to high-income male profiles increase significantly with local sex ratio, while those of the plain low-income women decrease.

Observation 1 confirms Hypothesis 1 and Hypothesis 2 for when the availability of high-income men increases. We address Hypothesis 1 and Hypothesis 2 for when the income of high-income men increases in Observation 2.

The opportunity costs from a possible drop in labor market participation after marriage and anticipated household income can be controlled for with the mean wages of men (*mean of men's income in a city*) and of women (*mean of women's income in a city*) in column (3) of Table 1. The significant coefficient for the standard deviation of women's income in a city (0.067) for female visitors suggests that women may also prefer higher income men when their own wages are more volatile. The insignificant coefficient for the standard deviation of men's income in a city (0.087) suggests that their prospective mates' wage dispersions may matter less to them. That is unsurprising, if these women are choosing the men for the men's higher level of income. Importantly, the increasing preference of high-income women for mate income remains even with all these controls.

The website allows for the reporting of only 9 income levels (<1, 1-2, 2-3, 3-5, 5-8, 8-10, 10-20, 20-50, and >50 in 1*k* CNY). Hence, we define *h*-, *m*-, and *l*-women by absolute cutoffs: *l*-women: <3*k*/month, *m*-women: 3-8*k*, *h*-women: >8*k* with the shares of *l*-, *m*-, and *h*-women being about 23, 62 and 15 percent of our visits, respectively. Our results are robust if we vary the cutoffs by +/- one level of income. The only case where our results become insignificant occurs when *h*-women are defined as those having incomes above 10*k*. However, only 4.5 percent of our female visitors fall into that category. This result is unsurprising if women have RDP for mate income and our *H*-men profiles have incomes in the range of 10-20*k*. We did not choose higher levels of male incomes to avoid outlier effects.

The local sex ratio for each city calculated from the 2010 Census includes migrants; thus, our findings may suffer from endogeneity because of unobserved factors in each city that affect both migration decisions and the preference for mate income, even after we control for various characteristics of individuals and cities. Therefore, we use the share of minorities in each city's population as an instrument for local sex ratio (Li and Zhang 2007; S.-J. Wei and Zhang 2011). The skewed distribution of local sex ratios (more men than women) in China is the result of traditional son preference, made more acute by the one child policy, under which people used ultrasound and other techniques to guarantee sons (Y. Chen, Li, and Meng 2013). Nonetheless, the one child policy was much less strictly applied for minorities than for the *Han* majority. Hence, if a higher proportion of minorities exists in a city, then local sex ratios should be less skewed (lower). However, the share of minorities should not affect people's preference for mate income, controlling for individual and other characteristics. A-Table 8 in Appendix 5 reports results of the instrumental variable ordered probit regression. As expected, in column (1), the share of minorities (*minority share*) is negatively correlated with local sex ratios. The second stage of the two-stage regression results in column (2) confirms the qualitative finding that the search intensity of high-income women still increases with local sex ratio relative to middle-income women in the same city (*h-women dummy*sex ratio*), even when we control for the women's beauty.

After demonstrating that only the search efforts of high-income women increase uniformly, irrespective of their beauty, when men became more plentiful, we now examine the effect of the changes in the incomes of the top-, middle- and bottom-1/3 income men (*H*-, *M*-, and *L*-men) in each city in Table 2 on the probability of these women visiting our high-income male profiles.[25] The insignificant coefficients for the mean income of men in column (1) reveal that women's average probability of visiting higher income male profiles seems not to be significantly affected by men's incomes. In column (2), we interact the three income levels with the *h*-women and the *m*-women dummies, but do not partition the women by their beauty, yet. Low-income women's search intensity (the benchmark) for high-income men decreases as the income of high-income men increases (-0.130). Again, this outcome could be due in part to the natural bias for findings of a negative coefficient for increases in the search intensity for high-income men, and therefore, false positive coefficient findings for increases in the search intensity for low-income men, as the income of high-income men increases. As with our finding for sex ratio in column (5) of Table 1, this decrease in low-income women's visits to the profiles of high-income men could also be because plain low-income women are switching from high- to low-income men as the competition for high-income men increases. Indeed, we will show evidence for this when we discuss column (4) of Table 2.

The probability of low-income women's visits to high-income male profiles increases

---

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

when the mean income of *M*-men increases (0.857) and decreases when the mean income of *L*-men increases (-1.036). The visits of medium-income women to high-income male profiles increase when the mean income of *L*-men increases (1.232). However, this increase is relative to the decrease in the probability of visits to higher income men among low-income women. The positive coefficient for high-income women (1.786) is partly an artifact of setting the negative coefficient of the low-income women as the benchmark and also partly because high-income women make a negligible share of their visits (3.5 percent) to low-income men. As expected, the visits of high-income women to high-income male profiles decrease when the mean income of *M*-men increases (-1.866). These findings anticipate our findings in Table 5, which show that the marriage probability of high-income women is negatively affected by the incomes of *H*-men, which should increase the competition these women face. The marriage probability of high-income women should be positively affected by an increase in the incomes of *M*-men, which should increase the supply of acceptable men, but not by the incomes of *L*-men, whom high-income women are averse to matching with, due to women's RDP.

In column (3) of Table 2, we include beauty and its interaction with sex ratio and the mean income of men with different income levels. We find the same pattern for increases in the income of high-income men in column (3) of Table 2 that we find for increases in sex ratio in column (5) of Table 1, when we divide the women of each income group by beauty and control for the previous results with sex ratio. Increases in the income of high-income men induce opposing reactions among the beautiful and the plain low-income women. When the income of high-income men increases, the significant negative coefficient of *mean income of H-men* (-0.366) indicates that plain low-income women are less likely to pursue (more likely to exit the market for) high-income men. In contrast, the coefficient for *mean income of H-men*beauty* (0.496) suggests that beautiful low-income women are more likely to enter the market for high-income men. However, the coefficient for entry is insignificant because it is known imprecisely, with a substantial standard error (0.316).

Similar to increases in sex ratio, the influence of beauty on women's response to the increase in the mean income of high-income men diminishes as women's income level rises, because of women's RDP. The significant positive coefficient for *mean income of*

*H-men*m-women dummy* (+0.404) in column (3) Table 2 indicates that the plain medium-income women are significantly 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* (-0.678) indicates that beautiful medium-income women are significantly less likely to enter than beautiful low-income women are when the income of high-income men increases. The significant positive coefficient for *mean income of H-men*h-women dummy* (0.802) indicates that plain high-income women are significantly 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 insignificant negative coefficient for *mean income of H-men*beauty*h-women dummy* (-1.071) indicates that beautiful high-income women are less likely to enter than beautiful low-income women. Again, this pattern of decreasing differentiation by beauty among women as women's income increase is expected if high-income women are more desperate (less willing to avail themselves of the option of low-income men) to match with a high-income man, due to their RDP.

[Insert Table 2 here.]

We calculate the marginal effects of sex ratio and men's income on high-income women's probability of visits based on the coefficients of the ordered logit regression in column (2) of Table 2, keeping all variables at their mean values. The 5.329 coefficient of *sex ratio*h-women dummy* indicates that a 10 percent increase in the sex ratio increases the probability of high-income women visiting high-income profiles by 9.09 percentage points, and decreases the probability of visiting middle- and low-income profiles by 7.84 and 1.25 percentage points, respectively, compared with l-women. Table 1 and A-Table 8 provide similar results. We also control for the effect of facial beauty in Table 2. The interaction of the *mean income of H-men*h-women dummy* remains significant (0.802) in column (3).

The 0.354 coefficient of *mean income of H-men*h-women dummy* in column (2) of Table 2 implies that a 10 percent increase in the mean income of *H*-men increases the probability of *h*-women visiting high-income male profiles by 15.94 percentage points, and decreases the probability of visiting middle- and low-income male profiles by 9.17 and 6.77 percentage points, respectively, compared with that of *l*-women.

These results in Table 2 are summarized in Observation 2.

***Observation 2.*** The visits of high-income women and beautiful low-income women to high-income male profiles increase with the mean income of high-income men, while those of the plain low-income women decrease.

Observation 2 confirms the part of Hypothesis 1 predicting that the search intensity of high-income women increases with the income of high-income men. This observation is also consistent with, although not significantly confirmatory of, the part of Hypothesis 2 predicting that beautiful low-income women also increase their search intensity when the income of high-income men increases. The loss of significance is due to the substantial standard error. One reason why the standard error for the search intensity of beautiful women for high-income men would rise is that the men themselves may exert greater effort in their search for a beautiful girlfriend/wife when their income increases,[26] making the women's own efforts for finding a high-income husband less necessary in many cases.

Indeed, Table 3 shows evidence consistent with this possibility; men's search intensity for beautiful women increases on the men's own income. Columns (1) and (3) report on how the beauty of the profile that men visit increases on the men's income. Column (1) uses the level of men's income. Column (2) uses a high-income men dummy. However, the search intensity of high-income men for beautiful women does not increase on sex ratio. 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) shows this by interacting sex ratio and the level of men's income, and column (3) shows this by interacting sex ratio and the high-income men dummy. This result (though not included in our hypotheses) is important for the interpretation of Observation 4 showing that the beauty of the wife of high-income men increases on sex ratio.

[Insert Table 3 here.]

***Observation 3.*** Richer men's probability of visits to beautiful female profiles is higher than poorer men's, but the richer men's visits do not increase significantly on sex ratio.

Hence, though richer men search more vigorously for beautiful women, that greater

---

[26] We thank Aloysius Siow for pointing out this possibility.

search intensity does not increase with the availability of men.

### *D. Beauty of the Wife of High-Income Men*

We now test for the expected consequences of low-income women's (beautiful or otherwise) increased search efforts for high-income men on the beauty of the wife of high-income men using the China Family Panel Studies (CFPS) 2010 baseline dataset. 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 for the whole population of China. We restricted the sample to married couples living in urban areas, both aged 20-45, comprising 2191 couples. We broaden the age range of couples here to increase the sample size. We dropped the couples in which the husband does not earn a positive income leaving us a final sample of 2147 couples for analysis. We use surveyor's 0-7 scale rating of the beauty of those they surveyed.

We follow Edlund et al. (2013) in constructing the sex ratio for the age range of 20-45 using data from the 2010 Census, which only reports ages at five-year intervals: 20-24, 25-29, 30-34, 35-39, and 40-44. We proxy the sex ratio faced by the husband and the wife at the time of marriage using a five-year window (adjacent two years above and below) with a two-year age gap between men and women.[27]

Consistent with the predictable consequences of our result that a greater number of women compete for high-income men when these men become more plentiful, the coefficient for *sex ratio*H-men dummy* in column (1) of Table 4 shows that the beauty of the wife of high-income men increases with sex ratio. Moreover, the beauty of the wife of high-income men is also higher (0.066, s.e. 0.68) than that of the wife of the medium-income men, which in turn is also higher (-1.45, s.e. 0.072) than that of the wife of the low-income men.

[Insert Table 4 here.]

[27] Thus, for example, for a 40-year-old woman, we use the ratio of the number of men age 40-44 to the number of women age 38-42. To impute the number of women in each age in the range 38-42, we take 1/5 of the number of women 35-39 for women in this age range, and likewise 1/5 of the number of women 40-44 for women in this age range. Then we sum up the number of women 38-42.

***Observation 4.*** The probability of high-income men marrying beautiful women increases with their income and on local sex ratio.

Thus, we confirm Hypothesis 3, which is consistent with high-income men enjoying a larger pool of more attractive women to choose from when the competition for them increases.

## IV. Census and Household Survey Data

### *A. Marriage Probability*

Here we test for the effects that we have found thus far of the accumulating entry of low-income women into the mating market for high-income men on the marriage probability of high-income women formulated in Hypothesis 4. It predicts that the marriage probability of high-income women decreases (relative to women of other income levels or absolutely) with the availability of high-income men, notwithstanding their increased search intensity, while that of low-income women would not be adversely affected, despite their search intensity not significantly increasing, in aggregate.

We use the 20 percent random sample of the China 2005 1 percent Population Survey, which contains micro level data.[28] The entire sample contains 2,585,481 individuals in 31 provinces in China.[29] We restrict the female sample to women aged 22-30 years to allow them to finish college and enter the marriage market, while not being too old for us to identify variations in their marital status. We restrict the male sample to those aged 22-35, years who are likely matches for these females. These males all have urban *hukou* and earn positive incomes. We exclude provinces with significant minority populations,[30] which can exhibit unique marriage matching traditions (Ji and Yeung 2014), obtaining a final sample of 29,593 women. We estimate the following logit model of the probability of being married for woman *i*

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

[28] Due to a change in the Government policy, such data were not released for the 2010 Census data.

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

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

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. *X* includes woman *i*'s *log* monthly wage, the local sex ratio (the *log* of the number of males over the number of females both aged 22-35 years in each city), and the mean income of *H*-, *M*-, and *L*-men (defined as the top-, middle-and bottom-thirds, respectively, of the income distribution of the male populations of each city). 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. Note also that though 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). We interact the dummy variables for the different categories of women with sex ratio and the mean income of *H*-, *M*-, and *L*-men. We use the mean income of men of different income categories within a city as the treatment variable because these are exogenous to women's individual incomes. The regression results are presented in Table 5.

The positive coefficient for sex ratio in columns (1) and (2) of Table 5 for the benchmark *l*-women and the insignificantly different *m*-women are consistent with the standard theory (Becker 1973) that women on average benefit from a higher sex ratio. However, the availability of men negatively affects the marriage probability of high-income women (*sex ratio*h-women dummy*) in columns (1) and (2) in which we vary the controls for men's and women's mean income across cities. Similar to column (4) of Table 2, the interaction of sex ratio and the high-income women dummy becomes insignificant (0.100) when we include the interaction between men's mean income and women's income levels in columns (3) in Table 5. This result may be due to the already mentioned strong positive correlation between sex ratio and men's mean income, which may allow a significant coefficient for sex ratio only when we do not interact the mean income of men of different income levels with women's income level, as in the column (2) of Table 2.

Column (3) shows that the marriage probability of low-income women increases weakly with the increases in the mean income of high-income men (0.342), which should expand the low-income women's options. The marriage probability of low-income women should decrease with the mean income of low-income men (-0.792), which should increase competition from medium-income women, under the RDP hypothesis. Importantly, the marriage probability of high-income women decreases both relative to low-income women (-1.663) and even absolutely with respect to a zero benchmark with increases in *H*-men's mean income. Both the absolute and the relative effects are the anticipated consequence of the entry of low-income women into the market for high-income men. The loss of significance of sex ratio for high-income women suggests that the increase in the proportion of high-income men drives the negative effect of the local sex ratio on the marriage probability of high-income women.

Unsurprisingly, women's marriage probability always decreases with their level of education. This is consistent with the possibility that women who have lower marriage market endowments (e.g., attractiveness to men) have better labor market endowments or work harder (Boulier and Rosenzweig 1984). However, this pattern is also consistent with our hypothesis that women's probability of marriage decreases on their own 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 due to the increase in competition from low-income women. To rule out this possibility, columns (4) and (5) additionally control for the effect of the relative supply of men with an education that is college or above to women with an education that is college or above (*Edu ratio*). The coefficient for *Edu ratio* and *Edu ratio*h-women dummy* are small and insignificant for both columns (4) and (5). Moreover, column (4) shows no change with respect to column (2) in either the magnitude or the significance of coefficient of the interaction between sex ratio and the high-income women dummy (-0.989). Column (5) similarly shows almost no change with respect to column (3) in either the magnitude or the significance of coefficient of the interaction between the mean income of high-income men and the high-income women dummy (-1.711).

The marriage probability of high-income women increases with the income of middle-income men (1.828) in column (3). This can be because the incomes of medium-income men (average bounds of 757-1122 CNY/month) overlap with those of the high-income women (average lower bound of 1019 CNY/month). When the income of medium-income men increases, some high-income women have a greater number of desirable men, and this effect may dominate the effect of a greater number of women desiring the same men. For this reason, we use the low-income women (average upper bound of 547 CNY/month) as the benchmark, because according to our theoretical framework, they are likely to have strictly more marital options than high-income women.

[Insert Table 5 here.]

The key result in Table 5 is the negative coefficient (-1.663) in column (3) for *mean income of H-men*h-women dummy.* This significant negative coefficient indicates that probability of marriage of high-income women decreases on the income of high-income men relative to the probability of marriage of low-income women, holding all variables constant. In terms of magnitude 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 3.54 percentage points compared to low-income women. Women's RDP for a mate with higher 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. Thus, women's RDP predicts that the probability of marriage of high-income women should be relatively lower than low-income women's when high-income men become more desirable. 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 would be of lower income than the high-income men. But, we in fact find a *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* (-1.663+0.342=-1.291).[31] Thus, we find evidence with Census data with married couples supporting prior work with online dating data that only women have a RDP for mate income (Hitsch, Hortaçsu, and Ariely 2010; Ong and Wang 2015).

The fact that the total effect is entirely from the significant interaction effect of the mean income of *H*-men with the *h*-women dummy rather than the insignificant level effect (1.926) of merely the *h*-women dummy further highlights the importance of escalating competition. This total effect translates into a decrease of 2.75 percentage points in their marriage probability for a 10 percent increase in the mean income of *H*-men. The effect is substantial in light of the large population residing in the cities of our study.

The finding that the probability of marriage of high-income women decreases at all with increases on men's incomes is remarkable because it contradicts 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 (-0.295) by the incomes of low-income men (*mean income of L-men*h-women dummy*).

***Observation 5.*** The marriage probability of high-income women decreases on the local sex ratio and the incomes of high-income men, while that of low-income women increases significantly on local sex ratio and weakly on the income of high-income men.

However, despite this counterintuitive result, consistent with standard theory, the average marriage probability at the bottom of Table 5 is always positive; on average, women benefit from higher sex ratio.

## V. Discussion and Conclusions

Beyond demonstrating women's RDP with a new set of online dating experiments

[31] The p-value for the F-test is less than 0.001.

across 15 major cities, we use variations in men's incomes and local sex ratio to explore the increasing burdens on high-income women from the escalating competition for even higher income men. When the local sex ratio or the income of high-income men increases so 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 (irrespective of their beauty) for high-income men (Observation 1 and Observation 2). In contrast, only plain low-income women measurably decrease their search intensity for high-income men, when the local sex ratio or the income of high-income men increases (Observation 1 and Observation 2). Consistent with high-income men passively enjoying a larger pool of attractive women to choose from when sex ratio increases, the beauty of the wife of high-income men increases on sex ratio (Observation 4), despite their search intensity for beautiful women not increasing (Observation 3). The ultimate consequence of the competitive entry of low-income women (regardless of beauty) into the mating market for high-income men is evident in the marriage probability of high-income women. Despite the greater search intensity of high-income women, their marriage probability decreases when there are more high-income men or when high-income men are richer. In contrast, the marriage probability of low-income women increases when there are more high-income men and increases weakly when high-income men are richer (Observation 5). These findings from an online dating field experiment, CFPS, and Census data attest to the novel effects on both men and women of the competition that escalates on the women's income, due to women's RDP for mate income. Our findings could help explain the increasing difficulties of high-income women in finding mates in China, despite the increasing scarcity of women.

As discussed above in Table 5, the fact that high-income women's marriage probability decreases absolutely as local sex ratio or the income of high-income men increases, and not simply relative to that of women of other income groups, is consistent with the dominance of women's RDP for a mate income in relation to a possible men's RDP for a lower income mate. If men prefer lower income women as wives or if that preference is to predominate in marriage matching, then we expect that the first-order effect of an increase in local sex ratio and the income of high-income men is to increase the marriage probability of high-income women. In the case of increases in the sex ratio, there are more high-income men, which given heterogeneity among the men in their aversion to high-income women, should increase the number of men who match with high-income women. In the case where the income of high-income men increases, the gender gap in income increases, which should decrease the aversion of the high-income men to marrying these high-income women. However, neither outcome was observed in our data. Thus, our findings along with prior work with online dating data (Hitsch, Hortaçsu, and Ariely 2010; Ong and Wang 2015) suggest that women's preferences may be driving the spousal aversion toward the husband earning less than the wife.

Women's RDP can be rationalized within the standard marriage market framework as the result of women's attempt to offset their labor market opportunity cost and maintain their standard of living after marriage even if they decrease their labor participation to specialize in household production. That women decrease labor market participation after marriage due to pregnancy (Lundberg and Rose 2000; Waldfogel 1997) and tend to marry men who earn more than they do is well recognized in the empirical literature on marriage (see the large sociology literature on hypergamy). There is evidence that women gain in terms of household income in the US, but men do not when they cohabitate or marry (Light 2004). However, to our knowledge, the theoretical literature on marriage matching has generally been gender neutral with regard to preferences for mate income.

Though Becker (2009) discusses the effect of women's labor market opportunity costs from adopting their traditional gender roles on their decision to marry or not, he does not discuss the possibility that women may seek out spouses who can cover the women's labor market opportunity costs (and therefore, support their habitual level of consumption). When preferences are symmetric across genders, either gender can specialize in household production. Labor market opportunity costs are symmetric *ex-ante* to the match. Who specializes in household production depends on who is in fact the lower income spouse. Following this line of reasoning, where couples decide on labor market participation *ex-post* of the match, there is no reason to expect that the competition that women face for mates will escalate with their income, as we find here.

However, diverging from the symmetric gender neutral preference assumption, recent research has taken seriously the possibility that one (e.g., Hersch (2013)) or both spouses prefer traditional gender roles. These studies presented empirical evidence that is difficult to explain otherwise. However, these studies have not drawn out in a theoretical framework the aggregate effect of such a preference on the direction and intensity of aggregate search efforts and marriage outcomes. In particular, they have not predicted that women will face competition for mates that escalates with their own income. This omission may also in part be due to the fact that the standard literature, until recently, has focused on a single measure of ability: income or education.

If men care only about mate income, for example, then high-income women can always "outbid" low-income women for high-income men. High-income women may still be able to outbid low-income women even if men also care about beauty. However, if these high-income women desire high-income men precisely to make up for their potential lost income after marriage, then high-income women may not be able to outbid attractive low-income women for high-income men, since in the long run, both types of women may not be working outside the home.

Recently, Becker type models have been extended to encompass multi-dimensional ability, e.g., education and BMI (Chiappori, Oreffice, and Quintana-Domeque 2012). Indeed, some part of the observations (that men's search intensity for beautiful women and the beauty of the men's wife increases with their income, and that the probability of marriage of high-income women decreases on the income of high-income men) can be explained within the standard framework (e.g., Choo and Siow (2006)), but with two-dimensional ability in our case: income and beauty. If men's marginal utility of spousal beauty increases with their own income, then plain high-income women may be less preferred than good-looking low-income women, when the income of the high-income men increases. However, despite this insight, and though men's increasing marginal utility for women's beauty also helps to further explain why the search intensity of beautiful low-income women's does not increase with the income of high-income men, how such assumptions on men's utility can explain our findings with sex ratio remains unclear. This assumption does not explain why high-income and beautiful low-income women's search intensity increases on sex ratio, nor why high-income women's probability of marriage decreases on sex ratio, nor why the beauty of the wife of high-income men increases on sex ratio. Nevertheless, our theoretical framework may still be a straightforward extension of the standard framework in Choo and Siow (2006) (tested in Siow (2015)) with a modification to the women's utility function. In particular, our results might be captured if women care about the spousal wage gap (i.e., the reference-dependent utility component) within a two-dimensional ability framework.[32] We leave the establishment of this connection for future work.

Although we use local sex ratio and changes in the income of high-income men as treatments to reveal the comparative statics of women's RDP for mate income, the effect we identify is hypothesized to be the inherent competitive consequence of women's RDP for mate income. Thus, we expect high-income women to experience greater difficulty in finding mates where high-income women themselves earn higher incomes (which limits the pool of men they may find acceptable) for standard opportunity costs reasons, and where high-income men have higher incomes (which increases the entry of lower income women into the market for such men), especially if sex ratios are not as high as in China. Indeed, the negative effects of the interaction of escalating competition with the desirability and availability of high-income men should apply wherever women have a RDP preference for mate income (e.g., in Russia, where our preliminary analysis indicates similar findings).

---

[32] We thank Aloysius Siow for pointing this out.

## References

Barberis, Nicholas C. 2013. "Thirty Years of Prospect Theory in Economics: A Review and Assessment." *Journal of Economic Perspectives* 27(1): 173–95.

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

Becker, Gary S. 2009. *A Treatise on the Family*. Harvard University.

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

Chan, Kam Wing, and Will Buckingham. 2008. "Is China Abolishing the Hukou System?" *The China Quarterly* 195(195): 582–606.

Chen, Hua, Sung H Ham, and Noah Lim. 2011. "Designing Multiperson Tournaments with Asymmetric Contestants: An Experimental Study." *Management Science* 57(5): 864–83.

Chen, Yuyu, Hongbing Li, and Lingsheng Meng. 2013. "Prenatal Sex Selection and Missing Girls in China: Evidence from the Diffusion of Diagnostic Ultrasound." *Journal of Human Resources* 48(1): 36–70.

Chen, Zhuoqiong Charlie, David Ong, and Roman M. Sheremeta. 2015. "The Gender Difference in the Value of Winning." *Economics Letters* 137(2015): 226–29.

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2012. "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market." *Journal of Political Economy* 120(4): 659–95.

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy* 114(1): 175–201.

Clark, Derek J., and Christian Riis. 1998a. "Influence and the Discretionary Allocation of Several Prizes." *European Journal of Political Economy* 14(4): 605–25.

Clark, Derek J, and Christian Riis. 1998b. "Competition over More Than One Prize." *American Economic Review* 88(1): 276–89.

Deryugina, T, and O Shurchkov. 2015. "Does Beauty Matter in Undergraduate Education?" *Economic Inquiry* 53(2): 940–61.

Du, Julan, Yongqin Wang, and Yan Zhang. 2015. "Sex Imbalance, Marital Matching and Intra-Household Bargaining: Evidence from China." *China Economic Review*

35(September): 197–218.

Edlund, Lena, Hongbing Li, Junjian Yi, and Junsen Zhang. 2013. "Sex Ratios and Crime: Evidence from China." *Review of Economics and Statistics* 95(5): 1520–34.

Ernst, Christiane, and Christian Thöni. 2013. "Bimodal Bidding in Experimental All-Pay Auctions." *Games* 4: 608–23.

Farber, Henry S. 2008. "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." *American Economic Review* 98(3): 1069–82.

Fincher, Leta H. 2012. "China's 'Leftover'women." *The New York Times*.

Gustafsson, Björn, and Shi Li. 2000. "Economic Transformation and the Gender Earnings Gap in Urban China." *Journal of Population Economics* 13(2): 305–29.

Hamermesh, Daniel S., and Jeff E. Biddle. 1994. "Beauty and the Labor Market." *American Economic Review* 84(5): 1174–94.

Hersch, Joni. 2013. "Opting out among Women with Elite Education." *Review of Economics of the Household* 11(4): 469–506.

Hitsch, Günter J., Ali Hortaçsu, and Dan Ariely. 2010. "What Makes You click?—Mate Preferences in Online Dating." *Quantitative Marketing and Economics* 8(4): 393–427.

Ji, Yingchun, and Wei-Jun J. Yeung. 2014. "Heterogeneity in Contemporary Chinese Marriage." *Journal of Family Issues* 35(12): 1662–82.

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

Kőszegi, Botond, and Matthew Rabin. 2012. "Reference-Dependent Consumption Plans." *American Economic Review* 99(3): 909–36.

Li, Hongbin, and Junsen Zhang. 2007. "Do High Birth Rates Hamper Economic Growth?" *Review of Economics and Statistics* 89: 110–17.

Light, Audrey. 2004. "Gender Differences in the Marriage and Cohabitation Income Premium." *Demography* 41(2): 263–84.

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

Neumark, David, and Andrew Postlewaite. 1998. "Relative Income Concerns and the Rise in Married Women's Employment." *Journal of Public Economics* 70(1): 157–

83.

Nielsen, Helena Skyt, and Michael Svarer. 2009. "Educational Homogamy: How Much Is Opportunities?" *Journal of Human Resources* 44(4): 1066–86.

Ong, David. 2014. "Height and Income Attraction: An Online Dating Field Experiment." *Working Paper*.

Ong, David, and Jue Wang. 2015. "Income Attraction: An Online Dating Field Experiment." *Journal of Economic Behavior and Organization* 111(March): 13–22.

Ong, David, Yu Yang, and Junsen Zhang. 2016. "Trophy Mates: The Competition with and for Beauty and Income in Chinese Cities." *Working Paper*.

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

Siow, Aloysius. 2015. "Testing Becker's Theory of Positive Assortative Matching." *Journal of Labor Economics* 33(2): 409–41.

Tversky, Amos, and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *Quarterly Journal of Economics* 106(4): 1039–61.

Waldfogel, Jane. 1997. "The Effect of Children on Women's Wages." *American Sociological Review* 62(2): 209–17.

Wei, Shang-Jin, and Xiaobo Zhang. 2011. "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China." *Journal of Political Economy* 119(3): 511–64.

Wei, Yan, Quanbao Jiang, and Stuart Basten. 2013. "Observing the Transformation of China's First Marriage Pattern through Net Nuptiality Tables : 1982-2010." *Finnish Yearbook of Population Research* (Croll 1981): 65–75.

You, Jing, Xuejie Yi, and Meng Chen. 2016. "Love , Life , and Leftover Ladies in Urban China Love , Life , and Leftover Ladies in Urban China." *Working Paper*.

Yuan, Yan, Zhao Rong, and Lihe Xu. 2012. "Sex Imbalance, Marriage Competition, and Entrepreneurship: Evidence from Rural China." *Working Paper*.

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

Zhu, Wei X., Li Lu, and Therese Hesketh. 2009. "China's Excess Males, Sex Selective Abortion, and One Child Policy: Analysis of Data from 2005 National Intercensus

Survey." *British Medical Journal* 338.

# Tables and Figures

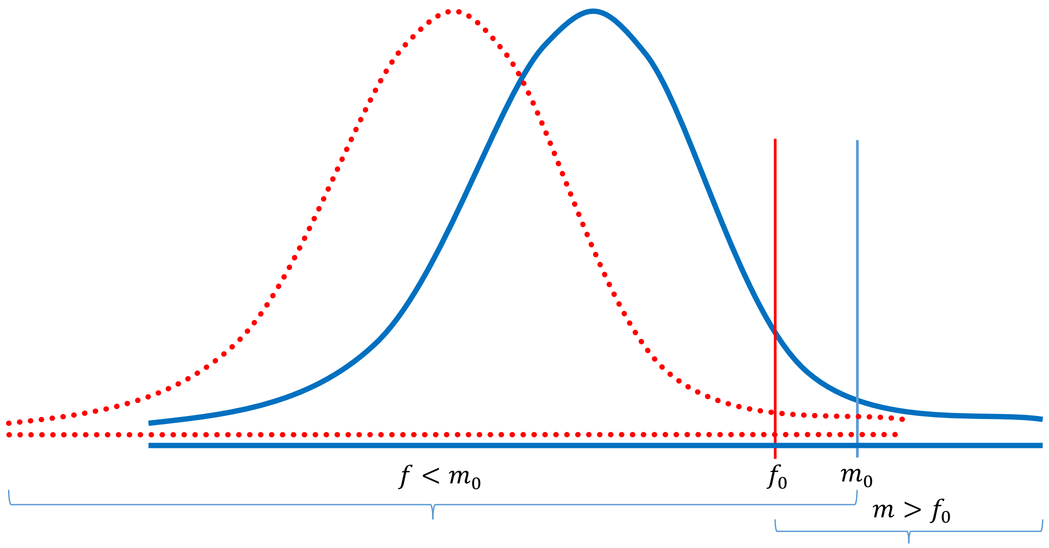



Figure 1: Men's and Women's Income Distributions and Women's RDP

Notes: The horizontal axis represents ranking by incomes of individuals of either gender (dotted red for women, solid blue for men), from lowest (LHS) to highest (RHS). The vertical axis represents the relative frequency of individuals with a specific level of income.


Figure 2: Share of Women's Visits by Male Profile Income Level Per Five-City Sex Ratio Group

Notes: 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. Then this number is 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 all 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 right-most graph shows a large gap between the visit rates for men who report an income 10-20k, which peaks at 16%, and those who report 8-10k, which peaks at roughly 9%, for all ages of men. This gap is much 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.

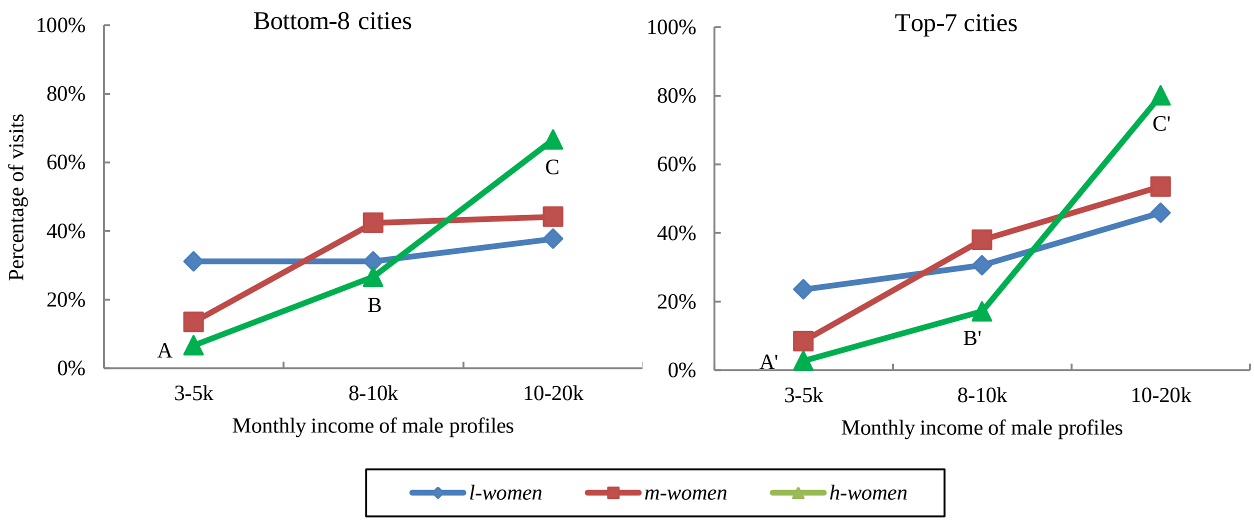



Figure 3: Share of Women's Visits to Male Profiles by Women Visitor's Income and Sex Ratio

*Notes:* We group women's visits into three income levels: <3, 3-8, and 8-20 (in 1*k* CNY), and label them as *l*-, *m-, and h-women*, respectively. These groups of visits are represented by three lines. We calculate the percentage of visits of each type of women to each type of male profiles. 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-8, indicating that women visited our high-income profiles more than our low-income profiles in the top- seven cities.

Table 1: Ordered Logit Regression of Women's Visits on Male Profile Income

| Dependent variable | Profile income (low (3-5*k*), middle (8-10*k*), high (10-20*k*)) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| *m*-women dummy | 0.603*** | 0.592*** | 0.420** | 0.412** | -0.632 |
| | (0.210) | (0.214) | (0.177) | (0.172) | (0.620) |
| *h*-women dummy | 1.780*** | 1.324*** | 1.143*** | 1.148*** | -0.471 |
| | (0.323) | (0.337) | (0.341) | (0.343) | (1.110) |
| **Sex ratio** | | **0.222** | **-0.130** | **-0.252** | **-16.657*** ** |
| | | (1.503) | (0.635) | (0.438) | (4.854) |
| Sex ratio**m*-women dummy | | 0.245 | 0.025 | -0.013 | 12.624** |
| | | (1.161) | (0.939) | (0.966) | (5.377) |
| **Sex ratio*h-women dummy** | | **8.829*** ** | **6.975*** ** | **6.783** ** | **28.759** ** |
| | | (3.617) | (3.228) | (3.326) | (13.979) |
| Beauty | | | | | -1.632 |
| | | | | | (1.051) |
| Beauty**m*-women dummy | | | | | 1.422 |
| | | | | | (1.177) |
| Beauty**h*-women dummy | | | | | 2.913 |
| | | | | | (2.074) |
| **Sex ratio*beauty** | | | | | **30.568*** ** |
| | | | | | (8.847) |
| Sex ratio*beauty**m*-women dummy | | | | | -21.593* |
| | | | | | (11.785) |
| Sex ratio*beauty**h*-women dummy | | | | | -48.232 |
| | | | | | (29.830) |
| Mean of men's income in a city | | | 0.059 | 0.078 | 0.147** |
| | | | (0.071) | (0.073) | (0.069) |
| Mean of women's income in a city | | | 0.005 | 0.013 | 0.004 |
| | | | (0.076) | (0.095) | (0.095) |
| s.d. of men's income in a city | | | 0.087 | 0.066 | 0.035 |
| | | | (0.087) | (0.098) | (0.088) |
| s.d. of women's income in a city | | | 0.067** | 0.061 | 0.042 |
| | | | (0.029) | (0.043) | (0.046) |
| GDP per capita | | | | 0.447 | 0.157 |
| | | | | (0.367) | (0.304) |
| Migration share | | | | -0.002 | -0.356 |
| | | | | (0.005) | (0.431) |
| Observations | 1,760 | 1,760 | 1,760 | 1,760 | 867 |
| Pseudo $R^2$ | 0.033 | 0.034 | 0.048 | 0.048 | 0.058 |

*Notes*: Data from the online dating experiment. Each observation is a visit (click) from a woman visitor. *m*-women dummy = 1 if woman's income is between 3k and 8k CNY/month. *h*-women dummy = 1 if woman's income is more than 8k CNY/month. The low-income women (omitted) is the benchmark. The local sex ratio is calculated using the 2010 Census, defined as the number of males/number of females (age 20-29 in 2010, or 24-33 at the time of experiment) in each city. Each woman's age, age$^2$, education years, and height are also controlled. Mean and s.d. (standard deviation) of men's and women's income are based on our online dating visitors age 22-35 and defined at the city level, and are in 1k CNY. GDP per capita is in *log* form. Migration share = residents without local *hukou*/total population in a city. Robust standard errors clustered at the city level in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

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

| Dependent variable | Profile income (low (3-5*k*), middle (8-10*k*), high (10-20*k*)) | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| *m-women* dummy | 0.457*** | -0.754 | -0.975 |
| | (0.176) | (0.590) | (0.827) |
| *h*-women dummy | 1.193*** | 1.316*** | 1.218 |
| | (0.336) | (0.785) | (1.888) |
| Beauty | | | 1.073 |
| | | | (1.370) |
| **Sex ratio** | -0.044 | -0.746 | **-7.834*** |
| | (0.998) | (0.892) | (4.113) |
| Sex ratio**m-women* dummy | -0.003 | -0.810 | 1.404 |
| | (0.862) | (1.188) | (2.107) |
| **Sex ratio*h-women dummy** | **6.989***** | **5.329***** | **3.184** |
| | (3.120) | (2.564) | (4.400) |
| **Sex ratio*beauty** | | | **14.516**** |
| | | | (6.386) |
| Mean income of *H*-men | -0.045 | **-0.130**** | **-0.366*** |
| | (0.054) | (0.059) | (0.188) |
| Mean income of *M*-men | 0.306 | 0.857*** | 2.060* |
| | (0.267) | (0.310) | (1.126) |
| Mean income of *L*-men | 0.045 | -1.036*** | -1.638 |
| | (0.253) | (0.302) | (1.362) |
| Mean income of *H*-men*beauty | | | **0.496** |
| | | | (0.316) |
| Mean income of *M*-men*beauty | | | -2.483 |
| | | | (1.910) |
| Mean income of *L*-men*beauty | | | 1.374 |
| | | | (2.891) |
| Mean income of *H*-men**m-women* dummy | | 0.068 | **0.404**** |
| | | (0.098) | (0.196) |
| Mean income of *M*-men**m-women* dummy | | -0.505 | -1.898 |
| | | (0.475) | (1.184) |
| Mean income of *L*-men**m-women* dummy | | 1.232*** | 1.784 |
| | | (0.288) | (1.417) |
| **Mean income of H-men*h-women dummy** | | **0.354***** | **0.802*** |
| | | (0.098) | (0.470) |
| Mean income of *M*-men**h*-women dummy | | -1.866*** | -4.939* |
| | | (0.503) | (2.558) |
| Mean income of *L*-men**h*-women dummy | | 1.786*** | 5.464** |
| | | (0.422) | (2.452) |
| Mean income of *H*-men*beauty**m-women* dummy | | | **-0.678**** |
| | | | (0.291) |
| Mean income of *M*-men*beauty**m-women* dummy | | | 2.852 |
| | | | (1.797) |
| Mean income of *L*-men*beauty**m-women* dummy | | | -1.326 |
| | | | (2.700) |
| **Mean income of H-men*beauty*h-women dummy** | | | **-1.071** |
| | | | (0.768) |
| Mean income of *M*-men*beauty**h*-women dummy | | | 7.200* |
| | | | (4.038) |

| | | | |
|---|---|---|---|
| Mean income of *L*-men* beauty**h*-women dummy | | | -8.518* |
| | | | (4.357) |
| Observations | 1760 | 1760 | 867 |
| Pseudo $R^2$ | 0.045 | 0.052 | 0.064 |

*Notes*: *m*-women dummy = 1 if woman's income is between 3k and 8k CNY/month. *h*-women dummy = 1 if woman's income is more than 8k CNY/month. The low-income women (omitted) is the benchmark. The local sex ratio is calculated using the 2010 Census, defined as the number of males/number of females (age 20-29 in 2010, or 24-33 at the time of experiment) in each city. *H*-, *M*-, and *L*-men = top-, middle- and bottom-1/3 men by monthly income in each city, respectively. The mean incomes are based on our online dating male visitors aged 22-35 and are in 1k CNY. Each woman's age, $age^2$, education years, and height are also controlled for. Robust standard errors clustered at the city level 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 | Female profile beauty ranking (within the range of 0-1) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Men's income** | **0.001**** | **0.001**** | | |
| | (0.000) | (0.000) | | |
| Sex ratio | | -0.096 | | -0.127 |
| | | (0.057) | | (0.078) |
| **Men's income*sex ratio** | | **-0.001** | | |
| | | (0.002) | | |
| *L*-men dummy | | | -0.006 | -0.008 |
| | | | (0.004) | (0.005) |
| ***H*-men dummy** | | | **0.013*** | **0.013*** |
| | | | (0.007) | (0.007) |
| Sex ratio**L*-men dummy | | | | 0.046 |
| | | | | (0.035) |
| **Sex ratio*H-men dummy** | | | | **0.001** |
| | | | | (0.044) |
| Constant | 0.647*** | 0.623*** | 0.662*** | 0.640*** |
| | (0.105) | (0.113) | (0.107) | (0.116) |
| Observations | 5,288 | 5,288 | 5,288 | 5,288 |
| $R^2$ | 0.022 | 0.025 | 0.022 | 0.024 |

*Notes*: Data from another experiment which was run simultaneously with 390 female profiles in the same 15 cities (Ong, Yang, and Zhang 2016). 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 to 8*k* CNY/month. The local sex ratio is calculated using the 2010 Census, defined as the number of males/number of females (age 20-29 in 2010, or 24-33 at the time of experiment) in each city. *L*-men = 1 if men's income is less than 5k CNY/month. *H*-men = 1 if men's income is more than 10k CNY/month. *M*-men (omitted) is the benchmark. Men's age, education year and height are also controlled. Robust standard errors clustered at the city level in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

Table 4: OLS Regression of Wife's Beauty on Husband's Income level

| Dependent variable | Wife's beauty |
|---|---|
| ***H*-men dummy** | **0.066** |
| | **(0.068)** |
| ***L*-men dummy** | **-0.145*** |
| | **(0.072)** |
| Sex ratio | 0.198 |
| | (0.468) |
| **Sex ratio****H***-men dummy** | **1.330*** |
| | (0.750) |
| Sex ratio**L*-men dummy | 0.265 |
| | (0.868) |
| Mean income of all men | -0.127 |
| | (0.373) |
| s.d. of income of all men | 0.047 |
| | (0.320) |
| Mean income of all women | 0.393 |
| | (0.291) |
| s.d. of income of all women | -0.035 |
| | (0.239) |
| Age | 0.000 |
| | (0.010) |
| Age of wife | -0.016 |
| | (0.010) |
| Income wife | -0.003 |
| | (0.008) |
| Education in years | 0.015* |
| | (0.008) |
| Education of wife in years | 0.044*** |
| | (0.010) |
| Constant | 3.177 |
| | (2.290) |
| Observations | 2,117 |
| R-squared | 0.133 |

*Notes*: Data are from CFPS 2010 restricted to married couples living in urban area and both are of age 20-45. *H*-, *M*-, and *L*-men = top-, middle- and bottom-1/3 men by monthly income in each city, respectively. All incomes are in *log* form. Robust standard errors clustered at province level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Logit Regression of Women's Marriage Probability

| Dependent variable | 1 = married, 0 = single | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **Sex ratio** | **1.092*** | **1.088*** | **0.300** | **1.073*** | **0.297** |
| | (0.593) | (0.588) | (0.552) | (0.593) | (0.569) |
| *m*-women dummy | -0.122 | -0.121 | 1.373 | -0.161 | 1.458 |
| | (0.099) | (0.099) | (1.273) | (0.148) | (1.299) |
| *h*-women dummy | 0.001 | 0.001 | 1.926 | -0.055 | 1.961 |
| | (0.097) | (0.097) | (1.254) | (0.157) | (1.253) |
| Sex ratio**m*-women dummy | -0.518 | -0.518 | 0.489 | -0.562 | 0.428 |
| | (0.526) | (0.526) | (0.451) | (0.522) | (0.458) |
| **Sex ratio*h-women dummy** | **-0.984*** | **-0.984*** | **0.100** | **-0.989*** | **0.067** |
| | (0.559) | (0.559) | (0.506) | (0.547) | (0.521) |
| **Edu ratio(men BA+/women BA+)** | | | | **0.052** | **0.124** |
| | | | | (0.131) | (0.116) |
| Edu ratio**m*-women dummy | | | | 0.061 | -0.059 |
| | | | | (0.132) | (0.108) |
| **Edu ratio*h-women dummy** | | | | **0.095** | **-0.054** |
| | | | | (0.143) | (0.113) |
| Mean income of *H*-men | | | 0.342 | | 0.382 |
| | | | (0.479) | | (0.484) |
| Mean income of *M*-men | | | -0.021 | | -0.147 |
| | | | (0.844) | | (0.832) |
| Mean income of *L*-men | | | -0.792 | | -0.804 |
| | | | (0.576) | | (0.589) |
| Mean income of *H*-men**m*-women dummy | | | -1.480*** | | -1.479*** |
| | | | (0.524) | | (0.532) |
| Mean income of *M*-men**m*-women dummy | | | 1.226 | | 1.246 |
| | | | (0.972) | | (0.981) |
| Mean income of *L*-men**m*-women dummy | | | 0.211 | | 0.179 |
| | | | (0.599) | | (0.604) |
| **Mean income of H-men*h-women dummy** | | | **-1.663*** | | **-1.711*** |
| | | | (0.557) | | (0.569) |
| Mean income of *M*-men**h*-women dummy | | | 1.828* | | 1.965** |
| | | | (0.998) | | (0.989) |
| Mean income of *L*-men**h*-women dummy | | | -0.295 | | -0.389 |
| | | | (0.637) | | (0.636) |
| Mean income of all men in a city | -1.042*** | -1.075*** | | -1.211*** | |
| | (0.134) | (0.412) | | (0.435) | |
| Mean income of all women in a city | | 0.035 | -0.433 | 0.197 | -0.322 |
| | | (0.423) | (0.368) | (0.447) | (0.383) |
| High | -0.497** | -0.490** | -0.498** | -0.515** | -0.490** |
| | (0.207) | (0.208) | (0.205) | (0.204) | (0.204) |
| Vocational | -0.833*** | -0.820*** | -0.833*** | -0.847*** | -0.828*** |
| | (0.205) | (0.206) | (0.202) | (0.204) | (0.202) |
| Bachelor | -1.300*** | -1.289*** | -1.301*** | -1.299*** | -1.288*** |
| | (0.198) | (0.198) | (0.195) | (0.196) | (0.195) |

| | | | | | |
|---|---|---|---|---|---|
| Constant | -20.657*** | -20.657*** | -22.602*** | -20.722*** | -22.639*** |
| | (2.526) | (2.527) | (2.923) | (2.527) | (2.937) |
| Observations | 29593 | 29593 | 29593 | 29,454 | 29,454 |
| Pseudo $R^2$ | 0.282 | 0.282 | 0.284 | 0.282 | 0.283 |
| Average effect of sex ratio | | | | | |
| | 1.092+ (-0.518)/3+ (-0.984)/3 | 1.088+ (-0.518)/3+ (-0.984)/3 | 0.300+ 0.489/3+ 0.100/3 | 1.073+ (-0.562)/3+ (-0.989)/3 | 0.297+ 0.428/3+ 0.067/3 |
| | **=0.591** | **=0.587** | **=0.496** | **=0.556** | **=0.462** |

*Notes*: Data are from China 2005 1 percent Population Survey restricted to women aged 22-30, with urban *hukou* and positive monthly income. Dependent variable = 1 if the woman is married, and 0 if single. *h*-, *m*-, and *l*-women, and *H*-, *M*-, and *L*-men = top-, middle- and bottom-1/3 men by monthly income in each city, respectively. All incomes are in *log* form. Each woman's age, age square, and province fixed effects are also controlled 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, the average effect will be a/3 + (a+b)/3 + (a+c)/3 = a + b/3 + c/3. Robust standard errors clustered at city level are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

# Appendices

## *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 | $z\theta - c$, $(1\text{-}z)\theta - c$ |
| | $Not\ Try\ (1-t)$ | 1, $\theta$ | 1, $\theta - c$ |

We want to show in this example that high-income women can be hurt when the prize they seek exclusively (high-income men) increases in value due to the increased competition (i.e., search effort/effort at attraction) from low-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), since 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.)

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, while 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 one low-income man, neither of whom have strategies. 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 $\theta$ 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 $\theta$ can put in $Effort$ at cost $c$ in getting him. She has no

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

interest in the low-income man. Hence, the outer nest (i.e., the low-income men) of the nested prize (i.e., 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) $\theta$ changes the competition between these two types of women.

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 both simplifying and raises the bar for our goal, which is to show that the high-income woman can be hurt when $\theta$ 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 $\theta > 1$, i.e., 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 that the payoff in each case is non negative. We first look for pure strategy Nash equilibria, which will be represented as ordered pairs of actions equivalent to a pair of degenerate probabilities over those actions. Then, we derived 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 $\theta$ 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 $\theta$ 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 $\theta$. When $\theta$ 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 $\theta$ to $\frac{(1-z)\theta-c}{1-z}=\theta-\frac{c}{1-z}$. If we increase $\theta$ 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 $(1-z)\theta-c$. Thus, the high-income woman's payoff increases with $\theta$ for some range, but then decreases discontinuously as $\theta$ 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: high- and low-income. We now 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 if shares of both the high- and the low-income women populations put in $Effort$ or $Try$, respectively. When the sex ratio increases, the high-income men are less scarce. The *ex-ante* effect of this decrease in the scarcity of high-income men prizes can be modelled now by substituting $\theta$ with $s\cdot\theta$, where $s\geq 0$ increases on the population of high-income men. It is obvious that we would find similar results as with an increase in $s$ that we find with an increase in $\theta$. Hence, the effect of an increase in sex ratio *ex-ante* to equilibrium will be similar to an increase in the income of high-income men. However, the effects of an increase in sex ratio in equilibrium, which takes into account potential shifts in competition between and within women of different income groups are more subtle, especially if men and women also differ by other characteristics, such as beauty.

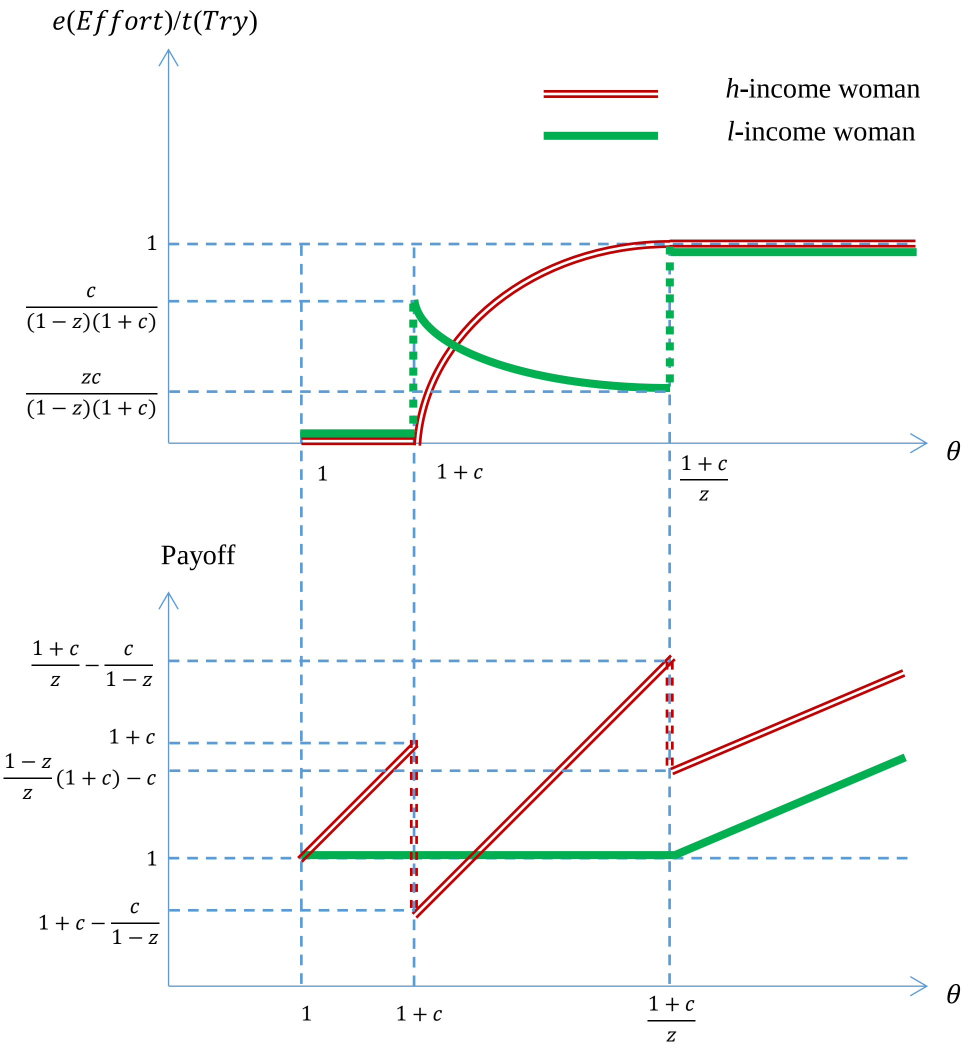
In real life, not only is the expected value of pursuing high-income men *ex-ante* to equilibrium higher when sex ratio increases, but the odds of getting a better looking low-income man is also higher, because there are a greater number of men for every woman. Thus, women in general not only have a better chance of getting a high-income men, but women of different levels of income and beauty have more scope to adopt heterogeneous strategies because there are more men. For example, if beautiful women of both high- and low-income levels pursue high-income men more because of increased availability, plain low-income women would face stiffer competition for these men. These plain low-income women may rather pursue better looking low-income men more -- despite the greater availability of high-income men. Thus, only certain subsets of women may actually enjoy the greater availability of high-income men in equilibrium, whereas others are crowded out. The effects of this extra level of heterogeneity according to beauty can be captured within our simple framework by the introduction of an extra coefficient $b > 1$ for $\theta$ for beautiful low-income women, and $b < 1$ for the plain low income women. When the low-income woman is beautiful and $b > 1$, her expected value of trying for the high-income man increases. She is more likely to enter. It is easy to see that this parameter $b$ will shift all thresholds to the left (e.g., $\frac{1}{b}$ instead of 1, $\frac{1+c}{b}$ instead of $1 + c$…etc in A-Table 2). Contrariwise if she is plain. We leave the detailed modelling of these very interesting potential equilibrium effects for future work, as our focus in this paper is empirical.[2] Our main goal here is to show that high-income women can be hurt by the increase in 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 $\theta$

| | $1 < \theta < 1 + c$ | $1 + \mathrm{c} < \theta < \frac{(1+c)}{z}$ | $\theta > \frac{(1+c)}{z}$ |
|---|---|---|---|
| $h$-woman | | | |
| $e(Effort)$ | 0 | $\frac{\theta-1-c}{(1-z)\theta}$ | 1 |
| payoff | $\theta$ | $\frac{(1-z)\theta-c}{1-z}$ | $(1-z)\theta - c$ |
| $l$-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 $\theta$, and the bottom row details that of $Try$ for the low ($l$)-income woman.

[2] We present evidence that lower income women adopt heterogeneous strategies according to their beauty when sex ratio (Observation 1) and the income of high-income men (Observation 2) increase.




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

*Notes*: $\theta$ is the value of the high-income man. The top panel illustrates the equilibrium strategy for the high ($h$)- and low ($l$)-income women, while 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 though the high-income woman is hurt 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 $\theta$ 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 $\theta > 1 + c$. However, the high-income woman has the advantage that she gets the high-income man by default and will not be challenged until the low-income woman is compensated for the loss of the low-income and search effort of $c$, which can be interpreted as search friction. However, as the value of $\theta$ 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. Cities Used in Online Dating Field Experiment*

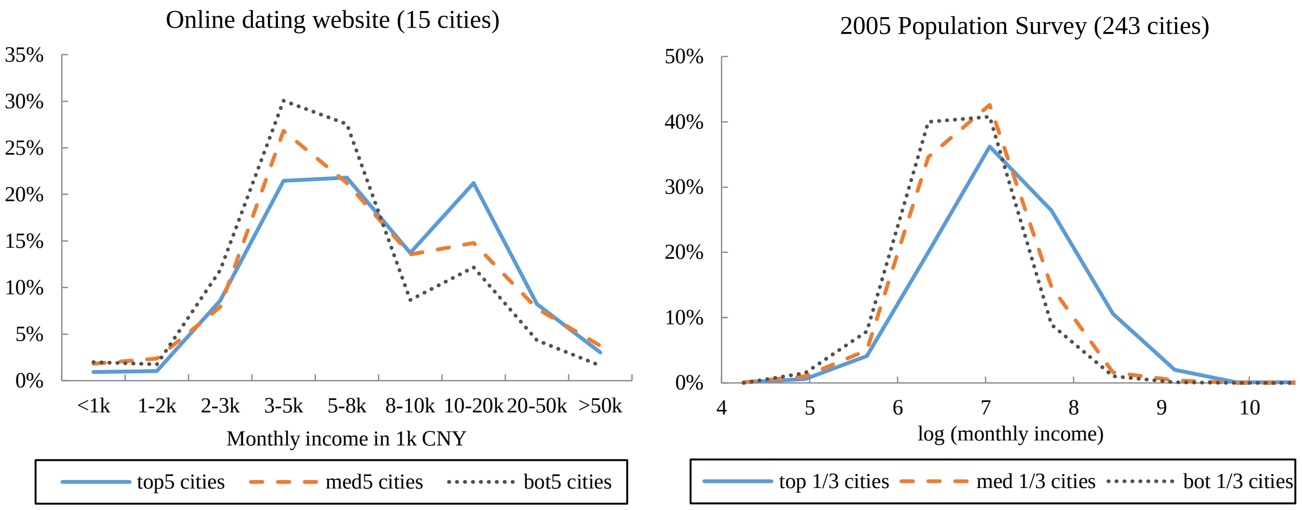
We started with 36 major cities (including all 31 provincial capitals and 5 vice-provincial level cities). We excluded 10 cities in minority provinces, and Ningbo, which is very close to Shanghai and Hangzhou, and Shenzhen which is too close to Hong Kong and may be affected by the Hong Kong marriage market. We also excluded three cities with age 20-29 and 25-34 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: GDP Per Capita, Income, and Local Sex Ratio in Cities Used in the Online Dating Experiment

| | City | 2013 GDP per capita | 2013 urban disposable income per capita | Age 20-29 sex ratio in 2010 | Age 25-34 sex ratio in 2010 |
|---|---|---|---|---|---|
| 1 | Tianjin | 101689 | 32658 | 1.262 | 1.302 |
| 2 | Guangzhou | 120516 | 42066 | 1.095 | 1.103 |
| 3 | Beijing | 92210 | 40321 | 1.070 | 1.076 |
| 4 | Nanjing | 98171 | 39881 | 1.059 | 1.049 |
| 5 | Xi'an | 57104 | 33100 | 1.053 | 1.020 |
| 6 | Shanghai | 90765 | 43851 | 1.050 | 1.074 |
| 7 | Xiamen | 81572 | 41360 | 1.044 | 1.084 |
| 8 | Shenyang | 88309 | 29074 | 1.038 | 1.018 |
| 9 | Hangzhou | 94791 | 39310 | 1.032 | 1.049 |
| 10 | Qingdao | 90746 | 35227 | 1.014 | 0.991 |
| 11 | Dalian | 110600 | 30238 | 1.008 | 0.994 |
| 12 | Zhengzhou | 68070 | 26615 | 1.006 | 1.030 |
| 13 | Chengdu | 63476 | 29968 | 0.990 | 1.014 |
| 14 | Jinan | 75254 | 35648 | 0.986 | 0.995 |
| 15 | Changsha | 99570 | 33662 | 0.975 | 0.978 |

*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.

## *Appendix 3. Income, Dispersion, and Local Sex Ratios*




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

Notes: The left panel exhibits the distribution of men's income for the 15 cities used in the online dating experiment divided into top-5 (top5 cities), middle-5 (mid5 cities) and bottom-5 (bot5 cities) 5-city groups in terms of the size of the local sex ratio. Local sex ratio is defined as the number of males/number of females ages 20-29 in the 2010 Census (which are ages 24-33 at the time of the experiments). The right panel exhibits the distribution of men's income in 243 cities ranked by local sex ratios, defined as the number of males/number of females ages 22-35 in the 2005 Census, and divided into top-1/3, medium-1/3 and bottom-1/3 local sex ratio city groups. The website only provides nine income categories, with the higher income categories encompassing a larger range of incomes (left panel), similar in scale to the *log* of income in 2005 Population Survey (right panel).

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 | 0.524*** | 0.524*** |
| | (0.178) | (0.175) |
| Men's income dispersion | | -0.003 |
| | | (0.158) |
| Province dummies | Y | Y |
| Constant | 6.742*** | 6.743*** |
| | (0.023) | (0.076) |
| Observations | 243 | 243 |
| R-squared | 0.535 | 0.535 |

Notes: Data are from the 2005 1 percent Population Survey. The sample is restricted to males and females aged 22-35 years and with urban *hukou* and a positive income. It excluded provinces with significant minority populations. 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 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 | 0.166** | 0.166* |
| | (0.079) | (0.085) |
| Men's mean income | | -0.001 |
| | | (0.045) |
| Province dummies | Y | Y |
| Constant | 0.506*** | 0.511* |
| | (0.011) | (0.304) |
| Observations | 243 | 243 |
| R-squared | 0.119 | 0.119 |

*Notes*: Data from the 2005 1 percent Population Survey. The sample is restricted to males and females aged 22-35 with urban *hukou* and a positive income. It excludes provinces with significant minority populations. 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.

## *Appendix 4. Summary Statistics of Visitors*

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 (Ong, Yang, and Zhang 2016) 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 height of 163 cm, and have a college degree, and 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 height of 175 cm, college degree, and income of 3-5, 8-10, and 10-20 k CNY/month. They are all unmarried with no children and are block randomly assigned to the 15 cities.

# *Appendix 5. Instrumental Variable Robustness Check*

A-Table 8: Ordered Probit Regression of Women's Visits with Instrument for Sex Ratio

| First-stage regression | | | Second-stage regression | | |
|---|---|---|---|---|---|
| Dependent variable | Sex ratio | | Dependent variable | Profile income (low (3-5*k*), middle (8-10*k*), high (10-20*k*)) | |
| | (1) | (2) | | (3) | (4) |
| Minority share | -0.017*** | -0.011*** | Sex ratio | 1.267 | -2.457 |
| | (0.001) | (0.003) | | (2.851) | (6.661) |
| *m*-women dummy | -0.005 | 0.014 | *m*-women dummy | 0.224* | -0.185 |
| | (0.004) | (0.016) | | (0.124) | (0.319) |
| *h*-women dummy | -0.005 | 0.029 | *h*-women dummy | 0.476* | -0.449 |
| | (0.007) | (0.026) | | (0.208) | (0.560) |
| Minority share*m*-women dummy | -0.001 | 0.002 | Sex ratio*m*-women dummy | 3.081 | 3.880 |
| | (0.001) | (0.005) | | (2.883) | (3.624) |
| Minority share*h*-women dummy | 0.001 | -0.003 | Sex **ratio*h*-women dummy** | **8.742*** ** | **20.383** ** |
| | (0.002) | (0.009) | | (2.961) | (8.454) |
| Beauty | | 0.059** | Beauty | | -1.059* |
| | | (0.025) | | | (0.543) |
| Beauty*m*-women dummy | | -0.054* | Beauty*m*-women dummy | | 0.921 |
| | | (0.031) | | | (0.642) |
| Beauty*h*-women dummy | | -0.067 | Beauty*h*-women dummy | | 1.669* |
| | | (0.048) | | | (0.976) |
| Minority*beauty | | -0.012* | Sex ratio*beauty | | 16.607*** |
| | | (0.006) | | | (4.675) |
| Minority *beauty*m*-women dummy | | -0.000 | Sex ratio*beauty*m*-women dummy | | -11.751* |
| | | (0.009) | | | (6.504) |
| Minority *beauty*h*-women dummy | | 0.011 | Sex ratio*beauty*h*-women dummy | | -25.380* |
| | | (0.016) | | | (15.195) |
| Constant | -0.680*** | -0.712*** | | | |
| | (0.083) | (0.116) | | | |
| Observations | 1,760 | 867 | Observations | 1,760 | 867 |
| F-statistic | 107.42*** | 36.67*** | | | |
| $R^2$ | 0.480 | 0.477 | | | |

*Notes*: Minority share = number of minorities/total population, in each city, using the 2010 Census. The local sex ratio, which is also calculated using the 2010 Census, is defined as the number of males/number of females (aged 20-29 in 2010, or 24-33 at the time of experiment) in each city. *m*-women dummy = 1 if female's income is between 3k and 8k CNY/month. *h*-women dummy = 1 if woman's income is more than 8k CNY/month. The low-income women (omitted) is the benchmark. Other control variables are the same as those in Column (4) of Table 1, and are omitted in this table. Robust standard errors clustered at the city level are in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.