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

Ong, David and Yang, Yu and Zhang, Junsen

Jinan University-University of Birmingham Joint Institute, University of Wisconsin, Chinese University of Hong Kong

2020

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

January 24, 2020

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

Reports of the difficulties of elite women in finding suitable mates have been increasing despite the growing availability and value of men in China. We rationalize this "leftover women" phenomenon within the directed/competitive search framework, which uniquely allows for equilibrium crowding out. Within this framework, we show that the leftover women phenomenon can be the result of women's aversion to men who have a lower income than themselves (hereafter, ALM) and the long-predicted complementarity between women's non-market traits (in particular, beauty) and male earnings. For high-income (h-)women, even when high-income (H-)men are more plentiful and richer, the direct effect of the greater number of desirable men can be overwhelmed by the indirect effect of competitive 'entry' by low-income (l-)women, particularly, the beautiful. We test for these competitive search effects using online dating field experimental, Census, and household survey data. Consistent with the competitive entry of l-women, when sex ratio and H-men's income increase, the search intensity of beautiful l-women for H-men increases. In response to this competitive entry, plain h-women, who are constrained by their ALM to search predominantly for H-men, also increase the search intensity. However, only their marriage probability decreases. Our evidence is consistent with intra-female competitive search for spouses who can cover the labor market opportunity cost of marriage and childbirth, which increases with a woman's income.

JEL Codes: C93, J01, J12

Keywords: directed search, marriage, sex ratio, online dating, aversion to lower income men, beauty

[*] We appreciate the constructive comments of the editor and two anonymous referees. We are grateful to Pierre-Andre Chiappori, Gordon Dahl, Jan Eeckhout, Joni Hersch, Ali Hortacsu, John Kennan, Sai Lan, Kevin Lang, Barton Lipman, Jessica Pan, Daniel Paserman, Kathleen Rybczynski, Sarah Rosenberg, Mark Rosenzweig, Aloysius Siow, Qing Wang, Shangjin Wei, Matthew Wiswall, Danyan Zha, Weilong Zhang, participants at the Econometrics Society North America, Asia, and China Summer Meetings 2019, the Society for the Economics of Households Meeting 2018, the Conference in Honor of Mark Rosenzweig 2017, Advances in Field Experiments Conference 2017, Asian and Australasian Society of Labor Economics Conference 2017, Tokyo Labor Economics Conference 2016, the NBER China Economy Working Group Conference 2016, the Chinese University of Hong Kong Economics Department Workshop on Family and Labor Economics 2016, the Econometrics Society Asia Meeting 2016, the Chinese Economic Association Meeting 2016, the University of Chicago Workshop on Inequality 2015, the Hitotsubashi Summer Institute on Labor Economics 2015, and seminar participants at the economics department of Harbin Institute of Technology, Hong Kong University, Lingnan University, McMaster University, National University of Singapore, Shanghai University of Finance and Economics, South China Normal University, Sun Yat-sen University, and Wilfrid Laurier University for their helpful comments.

[**] Junsen Zhang (corresponding author), jszhang@cuhk.edu.hk, School of Economics, Zhejiang University, Hangzhou, China; and Department of Economics, Chinese University of Hong Kong, Hong Kong, China. David Ong, dvdong@gmail.com, Jinan University-University of Birmingham Joint Institute, Jinan University, Guangzhou, China. Yu (Alan) Yang, alan.yang@wisc.edu, Department of Economics, University of Wisconsin-Madison, Madison, WI, USA.

I. Introduction

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

One might have supposed that the surviving women can only benefit from their scarcity. Indeed, this outcome is predicted by established economic theory; the short side of the mating market should enjoy a greater probability of matching and more surplus from their presumably greater bargaining power (Becker, 1973). Moreover, when women are scarce, men should compete harder to increase their mate value. Positive assortative matching predicts that high-income women, in particular, benefit when the income of high-income men increases. However, we show conceptually, through simulation, and empirically that if women are averse to a lower-income man (hereafter, ALM), then high-income women can be crowded out of the marriage market by lower income women when there are more men or when men are richer.

Such an aversion is suggested by Hitsch, Hortaçsu, and Ariely's (2010b) empirical study with US online dating data showing that women tend not to send first-contact emails to men who earn more than $20,000 less than themselves. Ong and Wang (2015) find that Chinese women have a similar aversion to a lower income men with an online dating field experiment with random assignment of income to male and female online dating profiles in China.[1] These revealed aversion findings with online daters are corroborated by Bertrand, Kamenica, and Pan's (2015) study showing that married couples "avoid" the situation where the husband earns less than the wife with US household survey data.[2] Men on the other hand are either indifferent to women's income (Goussé, Jacquemet, & Robin, 2018; Ong & Wang, 2015) or have a weaker preference than women for men's (Hitsch, Hortaçsu, & Ariely, 2010a). We outline a standard microeconomic basis for the gender difference in preference for mate income and women's

[1] See Appendix 1 for the graphs of their main results.

[2] See Appendix 2, where we replicate a key result in their paper of a discontinuous drop in the share of marriages when the wife's income exceeds that of her husband with Chinese Census data.

ALM in the literature review and conceptual framework sections. Here, we limit the discussion to an intuitive explanation.

The key insight for the comparative statics implications which we test in this study is that, with an ALM for mate income, increases in a woman's income reduce the pool of the men she finds attractive, while expanding the pool of other women who might prefer these same men. In the context of an ALM for mate income, the fierceness of the competition a woman faces for the men she most prefers 'escalates' as her income increases. The main focus of this present study is how this escalation in competition can, moreover, be exacerbated by increases in the income and availability of high-income men. Either may boost the expected return of pursuing such men: the former increases the value and the latter increases the probability of getting such a man.

The *direct effect* of increases in the income and the availability of high-income men benefit high-income women. In the case of increases in men's income, men are more desirable. In the case of increases in the local sex ratio (number of men/number of women in a city), more high-income men are available for each woman to desire. In either case, the higher *ex-ante* expected return for pursuing these high-income men may also increase the number of low-income women (particularly the beautiful among them) who might switch from pursuing low-income men to pursuing these high-income men as well.

Accordingly, an *indirect effect* of both increases in the income and availability of high-income men is the increased 'entry' of low-income women into the matching market for high-income men. As a result, high-income women, who are averse to matching with low-income men due to their ALM, are worse off in the marriage market. The indirect effect is likely to dominate the direct effect for high-income women, whereas the opposite is true for low-income women, who can also be satisfied with matching with low-income men, for whom they are not in competition with high-income women.

Consequently, high-income women can on balance be worse off in terms of marriage probability relative to low-income women or even in absolute terms when high-income men are even richer or more plentiful as a result of this 'competitive entry' of low-income women into the market for high-income men. Such may be the situation in China, where both the sex ratio and men's income compared with women's have been increasing dramatically (Ge & Yang, 2014). While our key finding that high-income women are less likely to match when there are more men to match with seems incompatible with Becker's (1973) theory of sex ratio, we show that it can, in actuality, be explained by the recent theory of directed or competitive search which builds on Becker's work. We discuss our results in terms of this literature at length in Section III Conceptual Framework.

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

Consistent with our competitive entry hypothesis, we show that the search intensity of the beautiful low-income women for high-income men increases with the local sex ratio and the income of high-income men. Consistent with substitution towards low-income men in the face of greater competition for high-income men, the search intensity of the plain-looking among the low-income women for high-income men also decreases with the local sex ratio and the income of high-income men. In contrast, consistent with a binding ALM, the search intensity of high-income women—in particular, the plain-looking among them—for the high-income men increases (though at a lower rate than the beautiful low-income women) with the local sex ratio and the income of these high-income men.

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

We also find evidence supporting prior findings that rising male wage inequality is a significant contributor to women delaying their marriage (Gould & Paserman, 2003).[3] To take this into account, we control for the mean and standard deviation of men's income in all regressions. Given these controls, we show with Census data that only high-income women experience a decreased marriage rate when high-income men's income increases, while low-income women's marriage rate increases. We provide an additional treatment variable, the local sex ratio, which we show contributes to the divergence of marriage rates among women of different income levels with Census data and among different income and beauty levels with household survey data. We, moreover, provide novel data on the search behavior of women and show that this search behavior varies with sex ratio according to the income and beauty level of the women, which given our ALM framework, predicts the divergence in marriage rates we find with Census and household survey data.

Our findings would be less surprising if the deterioration in the marital prospects of high-income women can be attributed to the cross-border migration of brides from low-income regions (Weiss, Yi, & Zhang, 2019). In fact, significant internal migration has occurred within China in recent years. However, because the local sex ratio includes the migrant population, the greater influx of low-income women migrants, which *reduces* the sex ratio, cannot be the driving factor for the association between a *higher* sex ratio and the decreased rate of marriage among high-income women. Moreover, the sex ratio is controlled for when we find that high-income women are adversely affected by increases in the income of high-income men.

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

II. Background and Related Literature

We build on Hitsch et al.'s (2010a) framework for analyzing first-contact email behavior to reveal user mate preferences. As they point out, the Gale-Shapley model, in particular, the deferred-acceptance algorithm approximates the behavior of online daters in their exchanges of signals of interest and emails if the search frictions they face are negligible. Hitsch et al. (2010a) show that preferences revealed in the online dating context are correlated with preferences revealed in actual marriages. We focus on the click-throughs of dating profiles on preliminary search results, which are prior to and necessary for making first-contact emails. Within the expected utility framework of our analysis, women's reaction to a given combination of match probability and match payoff associated with a male profile would vary with their preferences. Therefore, different reactions (here, click-throughs) among agents to a given profile would reflect both the variation in the probability of matching and the underlying preference of the woman for the match. For a fixed probability of matching, we can infer the woman's preferences using click-through rates.

Within our directed search framework, women's visit rate depends on both the women's preference for the men's income and the women's beliefs about their probability of success, i.e., on the expected value of the visit to the woman's income type, in equilibrium. The prior work of Ong and Wang shows that all types of women tend to visit scarcer high-income men (with whom they can expect a lower probability of success) with higher frequency within a given city with a fixed sex ratio. From this result, we infer that all income types of women prefer these high-income men and have an aversion to the more plentiful low-income men that is increasing on the women's income.

In this paper, we assume that women's ALM is constant across cities given that we controlled for individual characteristics of the female visitors and city-specific characteristics: the mean and variance of men's and women's income. Holding other factors (especially men's mean income) constant, we, therefore, can interpret that the increase in visits to the high-income male profiles when sex ratio increases as being due to the women's anticipation of an increased rate of success. In our experiment, we focus on the variation in the probability of individual women's visits to our male profiles across 15 cities due to changes in the sex ratio and the mean income of high-income men.

Our conceptual framework of escalating competition is similar to other theories in predicting that women's marriage rate decreases with their income (Isen & Stevenson, 2010) or educational attainment (Boulier & Rosenzweig, 1984). Recent research has focused on the potential effect of social norms in the US (Bertrand et al., 2015), Asia (excluding China) (Hwang, 2016), and other developed countries (Bertrand, Cortes, Olivetti, & Pan, 2018). We too do find women's marital prospects decrease with their educational attainment,[4] but not with income, once educational attainment is controlled for. However, entirely different from the prior literature, we examine the comparative statics effect of the local sex ratio and men's income on the search intensity and the probability of marriage of high- and low-income women.

Our study is closely related to but still differs significantly from the burgeoning literature on the effect of the local sex ratio on the competition for mates, particularly in China. Evidence exists for the expected increase in competition among men or their supporting families and the relaxation of competition among women. The rise in the local sex ratio predicts increases in men's level of entrepreneurship (Yuan, Rong, & Xu, 2012), men's work hours in dangerous and risky jobs (Wei & Zhang, 2011), male criminal activities (Edlund, Li, Yi, & Zhang, 2013), the savings of families with sons (Wei & Zhang, 2011), the time men spend on housework within households (Du, Wang, & Zhang, 2015), and women's participation in decision making (Edlund et al., 2013). The rise in the local sex ratio also decreases women's educational attainment and employment (Edlund et al., 2013). These studies on outcomes, which focus primarily on the behavior of men, confirm Becker's (1973) theory that the bargaining position of women improves as the sex ratio rises.

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

Women's ALM itself is a natural consequence of the classical assumption that women generally specialize in household production after marriage due to traditional gender roles

[4] Available on request.

(Becker, 1973).[5] This assumption is fleshed out in, for example, Caucutt et al. (2002), which posits that women take up the bulk of the burden of childcare and will marry only if the utility from pooled consumption and investment in children in the context of marriage is higher than their outside-option of being single (consume and/or raise children using their income).

With regard to women's labor market opportunity cost from specialization in household production, recent evidence supports the implication that Becker draws from the assumption that women specialize in household production after marriage, which is that marriage rates would increase with the gender gap in wages because the gap decreases women's labor market opportunity cost from specialization in household production (Autor, Dorn, & Hanson, 2018; Shenhav, 2020). Recent evidence also confirms Becker's prediction that marriage and childbirth decrease women's labor market participation based on the assumption that women tend to specialize more than men in household production after marriage. Women in the West (Lundberg & Rose, 2000) and in China (Feng, Hu, & Moffitt, 2017; Hare, 2016) often relinquish full-time work after marriage and childbirth. Among these women, the more highly educated are relatively more likely to "opt-out" completely (Hersch, 2013), possibly because their marriage to higher income men makes opting out more affordable (Bertrand, Goldin, & Katz, 2010).

Women's decreased labor market participation after marriage (Goldin, 2014) , and childbirth (Chung, Downs, Sandler, & Sienkiewicz, 2019) in the US Denmark (Kleven, Landais, & Søgaard, 2018; Lundborg, Plug, & Rasmussen, 2017), Norway (Andresen & Nix, 2019; Bütikofer, Jensen, & Salvanes, 2018), and China (Chen, Zhang, & Zhou, 2018), is the main source of gender differences in wages. Furthermore, women anticipate and report preferring (Parker & Wang, 2013) decreased labor market participation after marriage and childbirth in the US, even before they marry or graduate from college. Women also anticipate a relatively higher income husband (Wiswall & Zafar, 2018), who can potentially support their premarital standard of living. Correspondingly, men anticipate a lower income wife and no change in labor market participation after marriage.

Couple's anticipation of asymmetric labor market participation after marriage helps explain the asymmetry in preference for mate income revealed in search behavior found in Ong and Wang (2015). To the extent that a woman's income is expected to be forgone after marriage, her income will not benefit her potential husband. Men's apparent lack of income-based attraction to women as marriage partners may reflect their anticipation of the women's loss of

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

income. Given women's likely and anticipated reduction in income after marriage, it is natural for them to seek a husband whose income would substantially offset the opportunity cost of their *potentially* decreased labor market participation and specialization in household production to maintain their standard of living.[6] We provide a conceptual framework to understand and empirical evidence for the intra-gender competition between women resulting from their search for such husbands.[7] Moreover, we provide evidence consistent with the dominance of women's ALM for a higher income mate over a possible men's aversion to a higher income mate.

One of the main contributions of our study is to apply the directed search framework to marriage matching. Whereas there has been substantial theoretical work on directed search, the theoretical and empirical literature on directed search in the marriage market is sparse (Chade, Eeckhout, & Smith, 2017; Wright, Kircher, Julien, & Guerrieri, 2020). Among the exceptions are Xu and Yang (2016) which develops a search/matching model to analyze targeted search in a horizontally differentiated marriage market (where couples match by, e.g., personality). Arcidiacono, Beauchamp, and McElroy (2016) analyze the effect of sex ratio in a high school on sexual activity using a directed search framework. Few studies in the economics literature observe search behavior (Chiappori, Ong, Yang, & Zhang, 2017; Fisman, Iyengar, Kamenica, & Simonson, 2006; Hitsch et al., 2010b; Ong, 2016; Ong & Wang, 2015).

The closest to our work is Ong and Wang (2015), which identifies women's ALM for men of different income levels using visit shares of women of different income levels on the interaction of individual women visitor's income and individual visited male profile's income

[6] Such behavior is in-line with job search behavior in general which has been found to be reference-dependent on the prior job (Dellavigna, Lindner, Reizer, & Schmieder, 2017). In many cases, women literally also do search for jobs with less demanding more flexible hours, travel, and consequently lower pay after marriage and childbirth (Blau & Kahn, 2017). Men, in contrast, do not. This notion that a woman may search for a mate with a view to offsetting her opportunity cost is consistent with the long standing theory of habit formation and with the recent behavioral theory of reference-dependent preference in which the reference point is lagged consumption and women are loss averse (Kőszegi & Rabin, 2012).

Regardless of whether ALM has a standard microeconomic basis, we consider ALM as a primitive notion and focus on the potential over-entry/congestion effect of a targeted search for a mate based on income. In particular, we test for the possibility of crowding out of high-income women by the competitive entry of low-income women, when high-income men become more plentiful or richer.

[7] The intra-gender competition among women for eligible men has generally been a neglected topic. However, there is a rapidly expanding literature in anthropology on the competition between women (Stockley & Campbell, 2013). In economics, Ruffle and Shtudiner (2014) find a labor market beauty premium only for men and a beauty penalty towards women in a CV correspondence study in Israel. The beauty penalty was driven by in-house HR personnel, who are mostly young single women, rather than by HR firms. The authors infer that the bias against hiring more beautiful women by in-house HR personnel is a result of female sexual jealousy. While lab experiments find women generally less competitive than men in Niederle and Vesterlund (2007) type experiments, women have been found to be more competitive with each other than with men in these same experiments. Little work has been done on the possible competition between women for spouses who can compensate them for the labor market opportunity costs of motherhood. Some preliminary work along these lines shows that whereas male students' competitiveness predicts their own expected salary but not their expected spouse's, and female student's competitiveness predicts their expected spouse's salary, but not their own. This effect of women's competitiveness on their expected spouse's salary varies with the women's gender identity, defined as the degree with which they agree with the statement that "a man should put career first and a woman should put family first" (Jeon & Ong, 2018). Hwang (2016) shows that other East Asian societies (not including China) tend to exhibit high levels of agreement with such statements.

in five cities. While we also capture the interaction between individual female visitor and individual profile income in a parallel fashion, our main treatment variable is the interaction of local sex ratio and men's mean income with women's income dummies. We, moreover, allow for additional heterogeneity in the search behavior of the women by introducing interactions with the beauty ratings for our female visitors. These new interactions between aggregate variables and the beauty ratings allow us to capture the 'competitive entry' effect of beautiful low-income women into the marriage market for high-income men. It is this novel point of the competitive entry of low-income women into the market for high-income men that explains why high-income women's probability of marriage, particularly the plain-looking among them, decreases despite the increases in their visits to high-income male profiles.

III. Conceptual Framework

A frictionless matching framework, even if women are constrained by ALM to only consider men whose income is higher than her own, would not explain our main finding that the scarce side of the market is hurt by its scarcity, i.e., are less likely to find a match and get lower surplus from matching when there are more potential matches. The inadequacy of a frictionless matching framework is obvious when one thinks of such matching in terms of an optimal assignment problem. No matter what the marriage surplus is, increasing the number of potential matches cannot decrease the number of actual matches. In other words, in a frictionless setting, when one side of the market becomes more numerous, the possible matches of the other side must increase. The only way the matches of high-income women could decrease at the same time as the number of their potential matches increases is if the apparently necessary connection between the relative number of agents on different sides of the matching market and the match rate is broken.

We resort to the search literature for its greater level of flexibility, in particular, the directed search literature, which can explain our surprising finding that some part of the scarce side of the market is hurt by the increase in the availability or value of the other side. Directed search theory builds on the more traditional 'random' search theory, which itself introduces 'frictions' that allow for mismatches and non-matches in equilibrium to the matching theory framework of Becker. Agents in the random search framework choose to accept or reject potential matches arriving at random intervals based on a fixed exogenous distribution. These agents' main tradeoff is between shorter waiting time with low-value matches and longer waiting time with high-value matches. In our empirical setting, such frictions capture the infrequent meeting opportunities of women with high-income men in their daily life. Such frictions may be particularly severe for low-income women, because of sorting by educational and market institutions.

However, notwithstanding the increased realism of frictions of random search, marital search tends to be targeted in the sense that individuals use observable characteristics such as height, beauty, and signals of income to decide on meetings. Such targeted search is captured in the directed search framework where agents use observations of posted prices and wages to choose the set of potential meetings, e.g., jobs to apply to. Such targeting by price anticipates the match surplus and obviates the need for bargaining over the surplus split presumed in the random search framework. Because such surplus can be anticipated, applications of directed search models of marriage-matching tend to assume a non-transferable utility framework (Arcidiacono et al., 2016).

A key result in this literature, arising from the fact that agents choose the set of potential meeting partners simultaneously, is the possibility of *equilibrium* 'coordination failure' in the sense that multiple workers may show up for the same job, resulting in some workers remaining unemployed, while other vacancies remain unfilled because no workers showed up. This is because searchers' probability of matching are interdependent. In contrast to random search, the crucial tradeoff for agents in a directed search framework is not between match surplus and waiting time, based on an exogenous distribution of potential matches, but between match surplus and the *endogenous* match probability, which depends upon all other agents making the same tradeoff simultaneously in equilibrium (Chade et al., 2017; Eeckhout & Kircher, 2010; Wright et al., 2020). Within our context of changes in the local sex ratio or the mean income of high-income men, all women should re-adjust their tradeoff between high-probability/low-value matches and high-value/low-probability matches when the local sex ratio or men's income increases. Each woman's search exerts a negative externality on other women's probability of success in matching. Thus, to our knowledge, only a directed search framework can capture the crowding out phenomenon of plain-looking high-income women by beautiful low-income women that we observe.

While the theoretical literature at present offers little guidance on the equilibrium comparative static effects of the population distribution (e.g., the local sex ratio of a city in our study) for a particular submarket (e.g., the marriage market for high-income men), we argue that there are three factors which allow us to make a prediction about relative odds of marriage among different types of women within our setting.[8]

First, the increase in sex ratio or the mean income of high-income men increases the availability or the value of the scarce high-income men. As a mathematical necessity, these increases must ultimately decrease the severity of the tradeoff between match value and match probability faced by *some* types of women in *some* submarkets in equilibrium. Thus, in the case of the greater availability of such men from an increase in sex ratio or higher value of such men from an increase in the men's mean income, the expected value of searching for these men should increase for some if not all types of women in equilibrium.

A second factor which could be used for predicting the women's search intensity and relative odds of marriage is facial beauty. Becker (1973) predicts a tendency of beautiful women to match with high-income men. He posits that this is due to a complementarity between non-market productivity and money income.[9] The hypothesized complementarity between female beauty (proxied by youth) and men's income is empirically supported in the economics literature (Weiss et al., 2019). (See Appendix 3 for a formulation of the complementarity and Footnote 3 in Appendix 3 for the connection between female beauty and fertility, and male income and female fertility.) This complementarity between men with high-income and beautiful women may be manifested in behavior such as sexual attraction rather than in realized fertility. Due to the potential complementarity between male income and female beauty, the search behavior of beautiful women may increase more strongly than plain-looking women in response to changes in the availability or mean income of high-income men (because such women have more to gain from matching--particularly with high-income men).[10]

A third factor which could be used to predict the probability of marriage of different types of women is women's ALM. Women's ALM would predict that high-income women are less willing to make the tradeoff between match value and match probability than low-income women. In other words, while some low-income women may be willing to switch to searching for low-income men when the competition for high-income men increases, women's ALM suggests that high-income women would rather choose either to pursue high-income men with greater intensity or resort to their better singlehood option.

[8] We know of only one case where it has been worked out explicitly (Arcidiacono et al., 2016). Their theory does not apply to the case of heterogenous constraints on matching, as implied by our ALM assumption.

[9] See Section 4 of Part I of the Appendix of Becker (1973).

[10] This difference in search intensity that is increasing on the women's beauty is analogous to more productive firms in the labor market that are theorized to respond more to increases in the probability of matching because they have more to lose from an unfilled vacancy (Chade et al., 2017).

Hence, to summarize, we expect that the search intensity of beautiful low-income women for these men will increase as the availability or value of high-income men increases, and the marriage probability of plain-looking high-income women will decrease. We first provide a simple model with two types of men and women based on a standard directed search framework in Appendix 3 to show how low-income women switching from surer matches with low-income men to uncertain matches with high-income men, as the expected value of searching for such high-income men increases, may crowd out plain-looking high-income women from the marriage market. We then relax the discrete type assumption in a parametric simulation of the directed search and show that our predictions from the simple model still hold for a continuum of income and beauty types.

IV. Online Dating Field Experiment

Experimental Design

We outline the design of our experiment in following paragraphs. To preview the structure of our analysis in the rest of the paper, we use the experiment to identify the comparative statics effects of women's ALM on their search behavior as the sex ratio and men's mean income vary. This analysis of the experimental data forms the basis of our predictions for marital patterns observed in the Census and household survey data in subsequent sections.

Our field experiment extends Ong and Wang (2015) by testing for women's ALM 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 100 million members in 2016. The website we used (along with its competitors) advertises itself as a marriage-matching website for white-collar professionals between the ages of 25 and 45 years of age.

The users of this website can create a profile for free. The profile must include demographic (e.g., age and gender), socioeconomic (e.g., income), and physical characteristic (e.g., height) information, at least one photo, and a free-text personal statement. These requirements are standard to most online dating websites. Users may also add more information, and in particular, verifiable information to increase the "credibility" of their profile.[11] Users can browse, search,

[11] 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

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: an individual ID for the website, picture, nicknames, age, city, marital status, height, the first two lines of a free-text statement, and perhaps unique to China: income. Users can then click a link and 'visit' the full profile, where they can signal interest for free. Emails, however, require membership. The membership fee was 10 CNY/month at the time of the experiment when 1 USD was approximately 6 CNY. We recorded only visits.

We constructed our 450 male profiles on this website by collecting nicknames, pictures, and statements from real profiles from *another* website that would have automatically hidden them after a month of inactivity.[12] These profiles were posted for only 24 hours, after which, the accounts were closed. To further minimize any possibility of being recognized by acquaintances, we ensured that their picture was assigned to a province (city) that was different from their work area or birthplace.

We assigned 30 profiles of five ages: 25, 28, 31, 34, and 37; three incomes: 3-5, 8-10, and 10-20 (1*k* CNY) per month, which we will call low- (*L*-), middle- (*M*-), and high-(*H*-)income, respectively; and two replicas to each of the 15 major cities (see Appendix 4), resulting in 450 profile 'slots.' Then, we randomly assigned 450 pictures, nicknames, and personal statements to these 450 slots. For the profile's fixed traits, we gave all male profiles the height of 175cm. Birthdays were within eight days of each other and of the same zodiac sign. All of our profiles listed college education and the marital status of "single with no children" and "buy a house after marriage" (i.e., did not own a house).

Users can see our profiles' picture, nickname, age, city, marital status, height, income and the first few lines of a free-text statement in their default search results. They can then click a link and visit the full profile, which contained no additional information. For each of our profiles, we can see the profiles of the visitors by clicking their link in the history of visitors.

---

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.

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

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

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

See for example: http://humansubjects.stanford.edu/hrpp/Chapter9.html

The website records only visits to individual profiles once from any visitor even if they make multiple visits. Although visits across different profiles need not be from unique visitors, in our data, all visits were from unique visitors. This lack of repeated visits is expected because we took care to distribute our profiles between many other user profiles. Thus, to our knowledge, each of our data points is from a unique visitor. In any case, random assignment of characteristics to our profiles should rule out the individual idiosyncratic factors of our visitors as the primary driver of our findings.[13]

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

In total, our male profiles received 1,811 visits from women. 1,474 of these visits have photos. Among these, 1316 were of a quality useable for rating, e.g., of high enough resolution, had faces not obscured by sunglasses...etc. We had a random sample of two-thirds or 867 of these women visitor's photos rated for their beauty using a proprietary rating program accessible through a standard web browser. In the rating program, each female visitor's photo (i) is randomly matched with 10 other photos ($j \neq i$) from the pool of all photos. Each photo is selected with replacement from the pool of photos 20 times. Each photo was on average rated 200 times, which is approximately 10 times the frequency of other studies. A total of 692 Chinese raters (326 male) rated these 867 photos. The raters were graduate students from the Peking University HSBC Business School recruited through a mass email. We used two rounds for rating (one-third of photos in each round), because of our limited capacity to recruit raters during the first-round.[14]

[13] 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.

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

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

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

Description of the Data

The summary statistics of age, income, and education for each gender of our visitors are in A-Table 6 and A-Table 7 of Appendix 4. The women visitors range in age from 18-45. In 2005, the median age of first-marriage for women was 23 according to the 2005 1 Percent Population Survey (often called "the mini-Census"). 98.5 percent were married by the age of 30. The 2010 Census does not contain micro-level data of individual characteristics such as income or education. However, the CFPS data for 2010 show a similar pattern to the 2005 mini-Census. Accordingly, we find an abrupt decrease in the share of searches of women of all income levels, especially high-income women at age 30 in A-Figure 2.[15] We focus throughout the paper on women in the age range of 20 and 30 years old. We use the visits from women age 31 and 45 to check thatthe sex ratio we use does not affect the women outside of the age group that is the focus of our study.

We calculate the local sex ratio (the *log* of the number of men/number of women) using county-level data based on the full sample of the 2010 Census.[16] The local sex ratio for this

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

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

experiment is defined as the number of males between the ages of 22 and 32 years old over females between the ages of 20 and 30 years old at the time of experiment in 2014, proxied by males between the ages of 18 and 28 years old and females between the ages of 16 and 26 years old in the 2010 Census, which allows for the standard two-year age gap observed between married couples in China.

The individual characteristics of our women visitors that we collected include income, age, years of education, and height. The website allows for the reporting of only nine income levels (<1, 1-2, 2-3, 3-5, 5-8, 8-10, 10-20, 20-50, and >50 in 1*k* CNY). We define *h*-, *m*-, and *l*-women by absolute cut-offs: *l*-women, that is, <3*k*/month, *m*-women: 3-5*k*, *h*-women: >5*k*. We set these levels to approximate the bottom-, middle-and top-thirds of women, which are the groupings we use throughout the paper. For women between the ages of 20 and 30 years old, the shares of *l*-, *m*-, and *h*-women are approximately 25, 39 and 26 percent of our visits, respectively. The low-income level is the omitted benchmark in all regressions.[17]

In the following analysis, we show that the treatment effects of the sex ratio and men's income in our online dating sample are consistent with the treatment effects in the representative samples in the Census and CFPS datasets. For the comparison to be valid, we need to rule out that the consistency in treatment effects is due to a fluke of comparing disparate populations that happen to have similar treatment effects. We cannot directly compare the women visitors' characteristics to the women in the Census data because the 2010 Census does not contain individual-level income data.[18] (The 2015 mini- census is not currently available.) We also cannot make the comparison using the CFPS data because that dataset covers provinces and has minuscule samples for the cities used in our experiment. Instead, we compare the characteristics of our female visitors to the single women in the Urban Household Survey (UHS). The UHS is conducted by the National Bureau of Statistics and covers a representative sample of the urban population at the city-level in China. The UHS 2012 sample is the closest year to 2014 (the year of our experiment) available. It covers four of the cities used in our experiment, namely, Dalian, Shanghai, Guangzhou, and Chengdu. These cities are located in the Northern, Eastern, Southern, and Western China, respectively.

In comparison to single women in the UHS dataset, our female visitors are approximately four years older. 55 percent of single women in the UHS sample are younger than 26, whereas 26 percent of our female visitors are within that age range, with half concentrated in the age

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

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

range of 26 and 32. We find, similar to Hitsch et al. (2010a) in the US, that the website users in China are more educated (one year, in our case). In comparison to the UHS sample, 14 percent more of our women visitors have a college or higher level of educational attainment. (These distributional results are available on request.) Not surprisingly, given their older age and higher educational attainment, our women visitors also have a higher income than the general population.

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

[Insert Figure 1 here]

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

[Insert Table 1 here]

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

Graphical Analysis

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

[Insert Figure 2 here]

The upper panel shows the distribution of men's income on the website and in the surrounding city for the 15 cities used in the online dating experiment. The upper leftmost panel shows the distribution of reported incomes for our entire sample of 5,535 visits from men between the ages of 18 to 45 years old from the other experiment. The upper middle panel is restricted to the 3,520 visits from men between the ages of 25 to 37 years old, which is the age range that matches our male profiles. The upper right panel is restricted to the 2,832 visits from men between the ages of 22 to 32 years old, which is the age range of the men that we focus on throughout this paper. The lower left panel shows the income distribution of the men between the ages of 22 to 32 years old in the same 15 cities. The lower right panel shows that for the larger sample of 57 cities for which the 2005 mini-Census contained more than 300 respondents of either gender in the age range of 20 and 30 for women and 22 and 32 for men.[19] These 57 cities will be the focus of our analysis of women's marriage probabilities.

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

The upper panels show that our male visitors' income distribution of cities with high sex ratio (top-five-city group) is noticeably more right-skewed than that of those in the medium-, and bottom-five-city group for every age group. For the 2005 mini-Census sample in the lower panels, we observe a slightly greater right skew for the distribution of men's income for the middle-third sex ratio cities rather than the top-third. These panels suggest that male online daters in the 15 cities of our experiment and the men in the same cities for the 57 cities for

[19] The 2005 mini-Census is the last available with income data. We find qualitatively identical results when we use 250, 300, and 350 respondents.

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

which we have sufficient sample size are not poorer in higher sex ratio cities. The regression results for the 57 cities in A-Table 4 and A-Table 5 in Appendix 4 further suggest that the correlation between the local sex ratio and men's income is not negative. These results imply that increases in the sex ratio are not driven by a disproportionate increase in the share of low-income men compared with high-income men. Therefore,

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

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

[Insert Figure 3 here]

The graphs in Figure 3 exhibit visits by women to male profiles in each of the 15 cities of the experiment conditional on the local sex ratio, defined as the number of males between the ages of 22 and 32 years old over females between the ages of 20 and 30 years old at the time of experiment in 2014. These populations are proxied by males between the ages of 18 and 28 years old and females between the ages of 16 and 26 years old in the 2010 Census, which allows for the two-year age gap observed between married couples in China. The horizontal axis indicates the sex ratio of the city. 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 city divided by the visits to all our profiles over all male income types in the same city. The key finding is that the gap in the graphs of women's visits to middle- to high-income men and from low- to high-income men clearly increase on local sex ratios. The pattern suggests that the marginal impact of increases in men's income on the visits of women increases with the local sex ratio.

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

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

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

[Insert Figure 4 here]

These levels are labelled as *l-*, *m-,* and *h-women*, respectively and are represented by three lines. These three income levels for women are lower than the three incomes levels for our male profiles, because as in most countries, women in China typically earn less than men. The left (low sex ratio cities) and the right (high sex ratio cities) panels of Figure 4 show that women of all income levels visit high-income male profiles with greater probability. However, our focus here is not on the mate attraction effect of the absolute level of men's income on women's 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 reported incomes. Second, we show a kink in the graph of the high-income women at point B and B', which suggests that their visit rates to high-income male profiles (10-20*k*) increases significantly compared with that of middle-income male profiles (8-10*k*), that is, as the profile's income exceeds the women's average income (which is approximately 15*k*). Third, previewing the main findings in Table 2, Figure 4 displays a further counter-clockwise rotation from the left to the right panels (AB – BC and A'B' – B'C') from the cities with low sex ratio to those with high sex ratio for high-income women. For example, although a small percentage of *h*-women's visits were to male profiles with reported earnings of 3-5*k*/month in the bottom-eight sex ratio cities (point A), visits to such male profiles remain visibly lower to the point of being nearly zero in the top-seven cities (A'). Approximately 75 percent of *h*-women visits were to the 10-20*k* male profiles in bottom-eight sex ratio cities (C), but approximately 85 percent of their visits was to that type of profile in the top-seven cities (C'). These three levels of evidence suggest the increased search effort of high-income women, due to women's ALM, even before we impose controls econometrically.

Regression Analysis

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

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

where X includes *m-women dummy* (medium-income women), *h-women dummy* (high-income women), and *log* sex ratio (the *log* of the number of men/number of women—sex ratio from this point forward) and its interactions with the above two dummies, and individual and city characteristics.[21]

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

[Insert Table 2 here]

Table 2 displays the results of the ordered probit regression of women's visits as a function of their income and the local sex ratio for women in the prime age marriage market of 20-30 years of age. The positive term for the *h-women dummy* (0.502) in column (1) of Table 2 indicates that high-income women visit high-income male profiles more than low-income women do, thereby confirming our impression from Figure 4 and supporting previous findings (Ong & Wang, 2015). Column (1) of Table 2 demonstrates not only that the intercept for high-income women is higher than that of the benchmark low-income women, but also that the difference increases with the sex ratio (1.765). In apparent contradiction to our competitive entry hypothesis, the coefficient for *sex ratio* for the benchmark low-income women is small and statistically insignificant (0.432) in column (1) of Table 2.[22] However, the lack of significance can also be attributed to low-income women, who enjoy more outside-options among low- and medium-income men, behaving heterogeneously according to their beauty. In

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

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

particular, we show that the lack of significance is due to the less attractive among low-income women decreasing their search intensity for high-income men. This decrease exerts an offsetting effect on the increased search intensity of the beautiful low-income women for these same men.

We control for the beauty percentile ranking of female visitors in column (2) of Table 2 which we acquired for a random sample of two-thirds of these visits. Importantly for our previous finding of an insignificant increase in the search intensity of low-income women for higher income men, column (2) of Table 2 also reveals heterogeneity in the reactions of women according to their beauty rank and income level to increases in the sex ratio. The significantly negative coefficient for *sex ratio* (-6.219) indicates a pronounced decrease in the search intensity among of the benchmark plain-looking low-income women for high-income men when the sex ratio increases. By contrast, the highly significant positive coefficient for *sex ratio*beauty* (12.131) indicates a pronounced increase in the search intensity among the beautiful low-income women for high-income men when the sex ratio increases. Thus, increases in the sex ratio induce divergent reactions among beautiful and plain-looking low-income women, which helps to explain the apparent lack of reaction of low-income women in aggregate (when we do not disaggregate by beauty) in column (1) of Table 2.

The medium- and high-income women show diminishing contrasting reactions by their beauty rank as the sex ratio increases, due to their ALM. The significant positive coefficient (5.401) for the interaction between the sex ratio and the medium-income women dummy suggests that the plain-looking among them react less negatively than low-income women to the increase in the sex ratio. In contrast, the weakly negative coefficient (-8.339) for the interaction among sex ratio, beauty rank, and the medium-income women dummy suggests that the reaction of the beautiful among medium-income women to the sex ratio is less influenced by their beauty than those among low-income women. The positive and significant coefficient (10.879) for the interaction of the sex ratio and the high-income women dummy suggests that the plain-looking among the high-income women search more intensively for high-income men when these men are more plentiful. Similarly, in contrast, the negative coefficient (-19.219) for the high-income women suggests that their reaction to the sex ratio is less positively influenced by their beauty compared with that for low-income women.

This pattern of decreasing differentiation in search intensity for high-income men by women's beauty rank, as the women's income level increases, is expected because the lower the women's income level, the larger their set of options among low-income men, and therefore, the greater their latitude to avoid the increasing competition for high-income men.

Moreover, as indicated in column (3) of Table 2, these contrasting behaviors between beautiful and plain-looking women with different income levels become more significant for women between the ages of 20 and 30 years old and insignificant or reversed (-20.534) for women 31-45 in column (5), which suggests that these results are driven by women in the prime marriage market age in China. We largely confirm these contrasting behaviors for women between the ages of 20 and 30 years old with the Bartik (1991) style instrumental variable for the sex ratio in column (4) of Table 2,[23] though the levels of significance of some coefficients are reduced. See A-Table 8 for the first-stage results.

We calculate the marginal effects of the sex ratio on high-income women's probability of visits based on the coefficients of the ordered probit regression in column (3) of Table 2, for women 20-30 years of age, keeping all variables at their mean values. A 10 percent increase in the sex ratio decreases the probability of plain-looking low-income women's (ranked in the 25th percentile in beauty rank) visits to high-income male profiles by 24.2 percentage points.[24] In contrast, the same increase in the sex ratio increases the plain-looking high-income women's visits to high-income male profiles by 8.9 percentage points, which is approximately half the increase of the 16.7 percentage points of beautiful low-income women (75th percentile in beauty rank) to high-income men's profiles.

To summarize, the reaction to the increase in the sex ratio of the plain-looking among medium- and high-income women is less negative than that of the plain-looking among low-income women. The reaction of the beautiful among the medium- and high-income women is less positive than that of the beautiful low-income women. Thus, the greatest contrast between the behaviors of the high- and low-income women when the sex ratio increases is between the plain-looking high-income women and the plain-looking low-income women. The women's own beauty rank makes less of an impact on their search intensity for high-income men as the

---

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

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

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

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

women's income increases because their reluctance to substitute toward low-income men increases with the women's own income. This result is expected if high-income women are more determined (less willing to avail themselves of the option of low-income men) to match with a high-income man, when the competition for high-income men increases, due to women's ALM. These results are summarized by Observation 3.

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

We next examine the effect of the changes in the mean income of high-income men on the behavior of the women who visited our male profiles. We use the top-third of men by income who visited our female profiles in another experiment for our measure of high-income men.[25] These female profiles were in the same cities as this experiment on the visits of women to our male profiles. High-income men's mean income as compared to middle- and low-income men would determine the expected value of targeting the top-third of men within in a directed search framework. In equilibrium, women's beliefs about the mean income of such men are correct. Also, within a directed search framework, women's expected value of targeting such men takes into account the likely level of competition they face. We test for the effect of these men's mean income on the visit rates of our female visitors to our high-income male profiles in columns (1)-(3) of Table 3. We use the same abbreviations (*H*-, *M*-, and *L*-income) for these male visitors, as we do for the male profiles we created to attract female visitors, because these male visitors are of approximately the same income levels as our male profiles. *L*-men are the benchmark male visitors in our regression. We hold constant the income of the middle- and bottom-third income male visitors (*M*- and *L*-men, respectively). We omit discussion of the *M*-men because, being intermediate between *H*-men and *L*-men, the effect of the change in their mean income on high- and low-income women is ambiguous. We also omit the results for mean income of the benchmark *L*-men because they are almost always insignificant for *l*- and *m*-women. Moreover, less than 4 percent of *h*-women visit such men. We make these results available on request.

[Insert Table 3 here]

[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. For a given city, we divided the male visitors into top-, middle-, and bottom-thirds in terms of the income range they reported to the website, assigning them the midpoint of that range.

To begin, we control for the effect of the sex ratio in Table 3. As in column (3) of Table 2, the sex ratio remains negative but is now insignificant in column (1) of Table 3. Similar to increases in the sex ratio in column (3) of Table 2, the influence of beauty rank on women's response to the increase in the mean income of high-income men diminishes as women's income level rises in column (1) of Table 3, because of women's ALM. As with increases in the sex ratio in column (3) of Table 2, increases in the income of high-income men induce opposing reactions among the beautiful and the plain-looking low-income women in column (1) of Table 3.

When the income of high-income men increases, the significant negative coefficient of *mean income of H-men* (-0.483) indicates that plain-looking low-income women are less likely to pursue (more likely to exit the market for) high-income men. The insignificant coefficient for *mean income of H-men*beauty* (0.567) suggests that beautiful low-income women are weakly more likely to enter the market for high-income men when their mean income increases, and strongly (0.868) more likely to enter when we instrument the sex ratio in column (2) of Table 3. The significant positive coefficient for *mean income of H-men*m-women dummy* (0.615) in column (1) of Table 3 indicates that the plain-looking medium-income women are less likely than low-income women to exit (at least weakly more likely to enter) the market for high-income men.

In contrast, the significant negative coefficient for *mean income of H-men*beauty*m-women dummy* (-0.853) indicates that beautiful medium-income women are significantly less likely to enter than beautiful low-income women when the income of high-income men increases. The significant positive coefficient for *mean income of H-men*h-women dummy* (0.990) indicates that plain-looking high-income women are less likely to exit (at least weakly more likely to enter) the market for high-income men than low-income women. Similarly, in contrast, the significant negative coefficient for *mean income of H-men*beauty*h-women dummy* (-1.370) indicates that beautiful high-income women are less likely to enter than beautiful low-income women. These patterns are even more significant when we instrument the sex ratio in column (2) of Table 3, but they are completely absent for women age 31-45 in column (3). Again, as with the sex ratio results for women age 20-30 in column (3) of Table 2, this pattern of decreasing differentiation by beauty rank among women as women's income increases is expected if high-income women are more determined to match with a high-income man (and less willing to avail themselves of the option of low-income men), due to women's ALM.

Calculating the marginal effects based on column (1) of Table 3, a 10 percent increase in the mean income of high-income men decreases the probability of the plain-looking low-income women's visits to high-income male profiles by 13.3 percentage points. By contrast, that same increase in mean income increases the plain-looking high-income women's visits to high-income male profiles by 12 percentage points. Thus far, we find for plain-looking women a parallel change in behavior when either the sex ratio or high-income men's mean income increases. However, the parallel holds only weakly for beautiful low-income women. Beautiful low-income women's search intensity for high-income men increases insignificantly with the increase in men's mean income. We explain this apparent anomaly next.

The increase in the visit rate of plain-looking high-income women to our high-income male profiles when the mean income of other profiles rises is remarkable given that our high-income male profiles are then relatively less attractive. One would expect that the change in these women's rate of visits would be negative. Given this expectation of a negative coefficient, one should perhaps interpret the aforementioned lack of significance in the beautiful low-income women's visit rates to high-income male profiles as likely significantly positive. This weaker increase in the search intensity of beautiful low-income women for our high-income male profiles is also consistent with the possibility of these high-income men exerting greater effort in their search for a beautiful girlfriend/wife when their income increases, and thereby, obviating the need of these women to increase their own search intensity for high-income men. To preview, Table 4 presents evidence that is consistent with this possibility; men's search intensity for beautiful women increases with the men's income.

These results in column (1) of Table 3 are summarized in Observation 4.

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

Thus far, our results indicate that low-income women's entry into the market for high-income men when the sex ratio or high-income men's income increases increases with their (the women's) beauty rank. These results predict that the probability of high-income women (particularly the plain-looking) marrying decreases when either the sex ratio or high-income men's income increases. We test these predictions after testing for demand-side effects of men's search behavior. We divide men into rich and poor and show the interaction between these men's probability of visits to more beautiful women and the sex ratio in Table 4.

[Insert Table 4 here]

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

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

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

V. Results from Census and Household Survey Data

Marriage Probability

Here we test for the effects that we have found thus far of the accumulating entry of beautiful low-income women into the mating market for high-income men on the marriage probability of high-income women. These results predict that the marriage probability of high-income women, particularly the plain-looking among them, decreases with this competitive entry, whereas that of low-income women, including the plain-looking, is not adversely affected. We use the 20 percent random sample of the 2005 China mini-Census.[26] The entire sample contains 2,585,481 individuals in 31 provinces in China.[27] Again, we restrict the sample to women between the ages of 20 and 30 years old. Males earn a positive income, and both males and females have urban *hukou*.[28] We also excluded cities for which this mini-Census sample population of men and women in the prime age for the marriage market is below 300. Excluding the smaller cities makes our sample of cities here more similar to our sample of large cities in the online dating part of our study. Our findings are only slightly less significant if we

[26] 2005 is the latest year available which contains individual income data.

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

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

exclude cities with more than 250 or 350 of the men and women of interest. These results are available on request. Excluding cities in provinces with significant minority populations, which can exhibit unique marriage-matching traditions, we obtain a final sample of 20,929 women.[29]

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

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

where the dependent variable is the marital status of female i in city c. It equals 1 if the woman is married and 0 if she is single. The local sex ratio is again the *log* of the number of males 22 and 32 years old over the number of females between the ages of 20 and 30 years old in each city. The mean incomes of *H-, M-*, and *L*-men are defined as the top-, middle- and bottom-thirds, respectively, of the income distribution of the male populations of each city. With regard to these income levels, the average bounds across cities for men are 1,211-5,978 CNY/month for *H*-men, 757-1,123 CNY/month for *M*-men, and 194-702 CNY/month for *L*-men. The average bounds across cities for women are 1,019-3,197 CNY/month for *h*-women, 610-934 CNY/month for *m*-women, and 191-547 CNY/month for *l*-women. These ranges are not necessarily contiguous because the average bounds across cities are not the averages of bounds defined within each city.[30]

We interact the dummy variables for the different categories of women with the sex ratio and the mean income of *H-, M-*, and *L*-men. We use the mean income of men with different income categories within a city as the treatment variable because these are exogenous to women's individual income. Again, we omit discussion of the *M*-men because the effect of the change in their mean income on high- and low-income women will be ambiguous in our theoretical framework. We also omit *L*-men because less than 6 percent of *h*-women marry them. The regression results are presented in Table 5. Columns (1)-(3) present results for the same cities as in Table 2 for the online dating experiment. Columns (4)-(6) extend the sample to 57 cities for which we have a sample of more than 300 survey responses from each of men and women.

[Insert Table 5 here]

The weakly positive coefficient for the sex ratio in columns (1)-(6) of Table 5 for the benchmark *l*-women is consistent with our hypothesis that they may merely substitute towards high-income men when more of them are available. Controlling for educational attainment and the interaction of the mean income of men with different income levels with women's income

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

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

type in columns (2), (3), (5), and (6), high-income women are at least weakly more likely to be married than low-income women. Consistent with our findings of the competitive entry of beautiful low-income women into the market for high-income men when the sex ratio increases in column (3) in Table 2, the availability of men negatively affects the marriage probability of high-income women (*sex ratio*h-women dummy*) in all specifications in Table 5. In the case of column (1), where we restrict the sample to the 15 cities of the experiment and do not introduce interactions between the mean incomes of men with different income types, the larger standard error (2.362) makes the effect of the sex ratio insignificant. However, significance returns in column (2) when we introduce the interactions and in column (4) when we increase the sample to those 57 cities with more than 300 responses from each of men and women. We use the coefficient for sex ratio (-1.938) in column (5), where we also control for the interaction of the mean income of men with different income types of women, to calculate the marginal effect of the sex ratio evaluated at the mean values of all variables. A 10 percent increase in the sex ratio decreases the probability of high-income women marrying by 4.8 percentage points compared with low-income women, and by 1.2 percentage points in absolute terms, although the latter estimate suffers from a standard error that is too large to be significant.[31]

Moreover, consistent with our finding in column (3) in Table 2 of differentiation in search intensity by beauty rank among low-income women when the income of high-income men income increases, columns (2), (3), (5), and (6) in Table 5 show that the marriage probability of low-income women increases significantly with increases in the mean income of high-income men. This increase in high-income men's income should increase low-income women's probability of marriage because it strictly improves the low-income women's options among high-income men. Consistent with this competitive entry hypothesis, the marriage probability of high-income women decreases relative to low-income women and even absolutely with respect to a zero benchmark with increases in *H*-men's mean income.

Women's marriage probability can decrease with their level of educational attainment in the West (Isen & Stevenson, 2010) and in China. This is consistent with the possibility that women who have lower marriage market endowments (e.g., attractiveness to men) have better labor market endowments or work harder (Boulier & Rosenzweig, 1984). However, this pattern is also consistent with our hypothesis that women's probability of marriage decreases with their

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

opportunity costs, which may have a purely educational component. High-income women's probability of marriage may also decrease because the number of highly educated women rises faster than the number of highly educated men, rather than as a result of the increase in competition from low-income women. To control for this possibility, columns (3) and (6) additionally control for the effect of the relative supply of men with a college or above education to women with a college or above education (*Edu ratio*). The coefficient for *Edu ratio* and *Edu ratio*h-women dummy* is not significant for column (3) but is significant for column (6), which suggests that high-income women may benefit from more educated men. The interaction between sex ratio and the high-income women dummy and the interaction between the mean income of high-income men and the high-income women dummy are only slightly changed.

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

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

However, this negative *relative* effect of *H*-men's income on *h*-women's probability of marriage (relative to low-income women) is also consistent with a positive *total* effect of *H*-men's income on *h*-women's probability of marriage. In that case, the marriage probability of women of all income levels increases, but that of high-income women increases less than that of low-income women. Such a positive total effect is also consistent with a possible men's aversion for higher income mates. In the case of men's aversion to higher income mate, we expect that the first-order effect of an increase in the mean income of high-income men is to increase high-income women's marriage probability, because more of these women are of lower income than the high-income men. This is what we find in the 15 cities of the experiment. However, in the larger sample of 57 cities, we find a weakly *negative* total effect of men's mean income on high-income women's marriage probability, which is the sum of the interaction and the level of the *mean income of H-men* (-1.904+1.740=-0.164). Hence, our evidence does not support the men's aversion to higher income women hypothesis as the driver of our findings.[32]

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

We summarize our findings with the Census data as follows.

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

Somewhat supportive of standard theory, the average effect of the sex ratio on marriage probability at the bottom of Table 5 is positive for the larger sample of 57 cities with sufficient sample size, whereas it has a negative effect for the 15 cities of the online dating experiment. However, in both cases, the standard errors are too large for statistical significance. Hence, on average, women benefit somewhat from higher sex ratio. Our results suggest, however, that the weak effect of the sex ratio on the average woman belies its strong redistributive effects across women with different income levels.

We can gain further insight into which high-income women are losing ground to the competitive entry of beautiful low-income women with the China Family Panel Studies (CFPS) 2010 baseline dataset, which has beauty ratings of surveyed subjects.[33] The CFPS is a comprehensive survey of individual-, family-, and community-level data across China,

---

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

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

covering various aspects of economic and non-economic issues. It includes 16,000 households in 25 provinces and is representative of the whole population of China. We restrict the sample to married couples living in urban areas with a local *hukou*. We dropped the couples wherein the husband does not earn a positive income. Again, we constructed the sex ratio, which equals to the number of males between the ages 22 and 32 years of age over females between the ages of 20 and 30 years of age using data from the 2010 Census.[34] We restrict women to those between the ages of 20 and 30 years old, thereby leaving us a sample of size of 965. We use surveyor's 0 to 7 scale rating of the beauty of those they surveyed.[35]

We define the beautiful dummy = 1 for the top-20 percent of all women (rated 7 on the 1-7 scale). We again estimate the logit model of the probability of being married for woman i in Equation (2). Again, we define the mean income of *h-, m-*, and *l-wo*men to be the top-, middle- and bottom-third, respectively, of the income distribution of the female populations of each city. Table 6 shows the logit regression for the marriage probability for women with different income and beauty levels interacted with the local sex ratio and men's mean income in a city.

[Insert Table 6 here]

Table 6 shows that sex ratio has no effect on the marriage probability of either the plain-looking low-income women benchmark (0.945) or the beautiful-looking low-income women (-1.238). Table 6 also shows, however, that sex ratio does exert a significantly negative effect on plain-looking high-income women (-7.294), but a weakly positive effect on beautiful high-income women (9.817). In terms of marginal effects, the marriage probability of plain-looking high-income women is 10 percentage points lower than the plain-looking low-income women and actually declines 8.7 percentage points in absolute terms when the local sex ratio increases by 10 percent. This result suggests that the decline in the high-income women's probability of marriage as the sex ratio increases observed in the interaction of the sex ratio and the *h*-women dummy in all columns of Table 5 may be limited to the plain-looking among high-income women. Unfortunately, the small sample sizes we have for many provinces (e.g., less than half of the 29 have more than 30 data points for men's income) do not permit us to study the

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

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

potential linear relationship between women's marriage probability and the interaction of the mean income of men with different income types with CFPS data, as we did in Table 5 with the Census data.[36]

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

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

VI. Discussion and Conclusion

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

The consequence of this competitive entry of beautiful low-income women into the mating market for high-income men is evident in the marriage probability of low-and high-income women. Despite the greater search intensity of high-income women relative to low-income

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

women, their marriage probability decreases relative to low-income women's when there are more high-income men or when high-income men are richer (Observation 6). The marriage probability of low-income women increases weakly on the sex ratio and significantly on high-income men's mean income.

Analysis of the CPFS data reveals that it is specifically the plain-looking among high-income women whose marriage probability decreases relative to low-income women and even in absolute terms when the sex ratio increases (Observation 7). Moreover, the beauty of the wife of high-income men increases with the sex ratio. The fact that we did not find that high-income men's search for beautiful women increasing with the sex ratio (Observation 5) is consistent with the interpretation that the decrease in high-income women's marriageability when the sex ratio increases is the consequence of low-income women searching more intensively for high-income men when the sex ratio increases. Thus, taken as a whole, we provide substantial evidence for the intra-gender competition between women for high-income men that we hypothesized based on the prior evidence of ALM.

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

We emphasize that these high-income women are "leftover" in our framework not because they are undesired by men, as most men in China are in a desperate struggle to achieve the financial prerequisites for eligibility required by the ever-shrinking share of marriage-age women. Rather, high-income women remain single because their desire for the small share of men who are high-income (who could, therefore, compensate the women's substantial labor market opportunity costs from motherhood and specialization in household production) is unreciprocated by these men. This lack of reciprocity is to be expected given that these high-income men can compensate a large share of women, and therefore, can afford to be choosey. Thus, whereas men in China must be leftover by mathematical necessity, high-income women are only leftover by their *rational* choice, given their high opportunity costs of adopting traditional gender roles.

It is precisely the high-income women's better singlehood option which allows us to rationalize women's ALM and predict that the plain-looking among them are more likely to be single when the men they seek exclusively become more attractive to beautiful low-income women. Within the directed search framework, searchers tradeoff between match value and match probability in equilibrium. When there are more high-income men or when their mean income is higher, the severity of the tradeoff should be relaxed for at least some types of women. Given a complementarity between female beauty and male income, we would expect that beautiful women low-income women, who might have limited their search to low-income men, are most likely to benefit in redirecting their search to high-income men. Such substitution behavior by low-income women would crowd out the plain-looking high-income women from marrying the high-income men to their singlehood outside option.

Our finding that high-income women's marriage probability declines with sex ratio and the mean income of high-income men may be explained by an attendant higher reservation value for men's income on the part of high-income women, which leads to delay. Their reservation value may increase because of an increase in the dispersion in men's income, as was shown in the important work of Gould and Paserman (2003). However, we control for such dispersion in all our regressions. Thus, the decrease in marriage probability for high-income women, in particular, the plain-looking we find is on top of the delay due to dispersion in men's income.

Nevertheless, a greater number of men may increase the reservation value of the women and cause delay, even without increasing the dispersion of men's income, because the greater number of men increases the women's probability of matching. For this hypothesis to be true, the greater number of men would have to make only the plain-looking high-income women choosier since we find that only they suffer a decrease in marriage probability. However, we would expect that the greater number of men would also make the beautiful high-income women choosier. But, their marriage rate is not significantly affected by the local sex ratio. Furthermore, our results are not consistent with the alternative hypothesis that they are driven by men's aversion to higher income mate instead of women's aversion to a lower income mate. If that were the case, high-income women's probability of marriage should increase as high-income men become richer, because these richer men would find a greater number of high-income women earning less than them.

According to our hypothesis, women's ALM is driven by their labor market opportunity cost of motherhood. In that case, ALM should weaken with decreases in foregone earnings in terms of lifetime income due to delayed childbirth (Leung, Groes, & Santaeulalia-Llopis, 2016). Moreover, high-income women who are willing to delay may have lower fertility intent, and therefore, lower ALM as female fertility declines steadily after age 30, becoming more rapid after age 35.[37] Indeed, Brandt et al. (2018) find with Census data (which do not contain information on beauty) that while highly educated women's marriage rate catches up after age of 35, the realized fertility of women who marry after age 30, which has consistently been lower than women who marry earlier, has been declining steadily. The delay in marriage and decline in fertility is expected if these highly educated women faced increasing competition for high-income men as sex ratio and men's income relative to women's has increased, as they have in China.

Our findings based on online dating field experimental, household survey, and Census data demonstrate the novel, and to our knowledge, unexplored comparative statics effects of women's ALM for mate income. We suggest that these comparative static effects are the consequence of women's attempt to cover their labor market opportunity cost of household specialization after marriage and childbirth (which increases with their income) with the shared income of the man whom they marry.

[37] See the Clinical Guidance at the American College of Obstetrics and Gynaecology: https://www.acog.org/Clinical-Guidance-and-Publications/Committee-Opinions/Committee-on-Gynecologic-Practice/Female-Age-Related-Fertility-Decline?IsMobileSet=false

VII. References

Abramitzky, Ran, Delavande, Adeline, & Vasconcelos, Luis. (2011). Marrying Up: The Role of Sex Ratio in Assortative Matching. *American Economic Journal: Applied Economics*, *3*(July), 124–157.

Andresen, Martin Eckho, & Nix, Emily. (2019). What Causes the Child Penalty? Evidence from Same-Sex Couples and Policy Reforms. *Working Paper*.

Arcidiacono, Peter, Beauchamp, Andrew, & McElroy, Marjorie. (2016). Terms of Endearment: An Equilibrium Model of Sex and Matching. *Quantitative Economics*, *7*(1), 117–156.

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

Bartik, Timothy J. (1991). Boon or Boondoggle? The Debate Over State and Local Economic Development Policies. In *Who Benefits from State and Local Economic Development Policies?* (Vol. 45, pp. 1–16). Upjohn Institute for Employment Research.

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

Bertrand, Marianne, Cortes, Patricia, Olivetti, Claudia, & Pan, Jessica. (2018). Social Norms, Labor Market Opportunities, and the Marriage Market Penalty for Skilled Women. *Working Paper*.

Bertrand, Marianne, Goldin, Claudia, & Katz, Lawrence F. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, *2*(3), 228–255.

Bertrand, Marianne, Kamenica, Emir, & Pan, Jessica. (2015). Gender Identity and Relative Income Within Households. *Quarterly Journal of Economics*, *130*(2), 571–614.

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

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

Brandt, Loren, Turner, Laura, Li, Hongbin, & Zou, Iaqi. (2018). Are China's "Leftover Women" Really Leftover?: An Investigation of Marriage Market Penalties in Modern-day China, 1–53.

Bruch, Elizabeth E., & Newman, M. E. J. (2018). Aspirational Pursuit of Mates in Online

Dating Markets. *Science Advances*, *4*(8), eaap9815.

Burdett, Kenneth, Shi, Shouyong, & Wright, Randall. (2001). Pricing and Matching with Frictions. *Journal of Political Economy*, *109*(5), 1060–1085.

Bütikofer, Aline, Jensen, Sissel, & Salvanes, Kjell G. (2018). The Role of Parenthood on the Gender Gap Among Top Earners. *European Economic Review*, *109*, 103–123.

Caucutt, Elizabeth M., Guner, Nezih, & Knowles, John. (2002). Why do Women Wait? Matching, Wage Inequality, and the Incentives for Fertility Delay. *Review of Economic Dynamics*.

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

Chen, Yi, Zhang, Hong, & Zhou, Li-An. (2018). Motherhood and Gender Wage Differentials Within the Firm: Evidence from China. *SSRN Electronic Journal*.

Chiappori, Pierre-André, Ong, David, Yang, Yu, & Zhang, Junsen. (2017). Marrying Up: Trading off Spousal Income and Height. *Working Paper*.

Chung, Yoon Kyung, Downs, Barbara, Sandler, Danielle, & Sienkiewicz, Robert. (2019). The Parental Gender Earnings Gap in the United States. *Working Paper, Center for Economic Studies, U.S. Census Bureau*.

Dellavigna, Stefano, Lindner, Attila, Reizer, Balázs, & Schmieder, Johannes F. (2017). Reference-Dependent Job Search: Evidence from Hungary. *Quarterly Journal of Economics*, *132*(4), 1969–2018.

Du, Julan, Wang, Yongqin, & Zhang, Yan. (2015). Sex Imbalance, Marital Matching and Intra-household Bargaining: Evidence from China. *China Economic Review*, *35*(September), 197–218.

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

Eeckhout, Jan, & Kircher, Philipp. (2010). Sorting and Decentralized Price Competition. *Econometrica*, *78*(2), 539–574.

Feng, Shuaizhang, Hu, Yingyao, & Moffitt, Robert. (2017). Long Run Trends in Unemployment and Labor Force Participation in Urban China. *Journal of Comparative Economics*, *45*(2), 304–324.

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

Fisman, Raymond, Iyengar, Sheena S., Kamenica, Emir, & Simonson, Itamar. (2006). Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment. *Quarterly Journal of Economics*, *121*(2), 673–697.

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

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

Gould, Eric D., & Paserman, M. Daniele. (2003). Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates. *Journal of Urban Economics*, *53*(2), 257–281.

Goussé, Marion, Jacquemet, Nicolas, & Robin, Jean-marc. (2018). Marriage, Labor Supply, and Home Production. *Econometrica*, *85*(6), 1873–1919.

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

Hare, Denise. (2016). What Accounts for the Decline in Labor Force Participation Among Married Women in Urban China, 1991-2011? *China Economic Review*, *38*, 251–266.

Harsanyi, John C. (1973). Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points. *International Journal of Game Theory*, *2*(1), 1–23.

Hersch, Joni. (2013). Opting Out Among Women with Elite Education. *Review of Economics of the Household*, *11*(4), 469–506.

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

Hitsch, Günter J., Hortaçsu, Ali, & Ariely, Dan. (2010b). What Makes You Click?-Mate Preferences in Online Dating. *Quantitative Marketing and Economics*, *8*(4), 393–427.

Hopcroft, Rosemary L. (2006). Sex, Status, and Reproductive Success in the Contemporary United States. *Evolution and Human Behavior*, *27*(2), 104–120.

Hwang, Jisoo. (2016). Housewife, "Gold Miss," and Equal: The Evolution of Educated Women's Role in Asia and the US. *Journal of Population Economics*, *29*(2), 529–570.

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

Jeon, Gahye, & Ong, David. (2018). Is Female Competitiveness Influenced by Gender Identity? *Working Paper*.

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

Jokela, Markus. (2009). Physical Attractiveness and Reproductive Success in Humans: Evidence from the Late 20th century United States. *Evolution and Human Behavior*, *30*(5), 342–350.

Kleven, Henrik Jacobsen, Landais, Camille, & Søgaard, Jakob Egholt. (2018). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics, Forthcoming*.

Knight, John, Deng, Quheng, & Li, Shi. (2017). China's Expansion of Higher Education: The Labour Market Consequences of a Supply Shock. *China Economic Review*, *43*, 127–141.

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

Lang, Kevin, Manove, Michael, & Dickens, William T. (2005). Racial Discrimination in Labor Markets with Posted Wage Offers. *American Economic Review*, *95*(4), 1327–1340.

Leung, Man Yee Mallory, Groes, Fane, & Santaeulalia-Llopis, Raul. (2016). The Relationship between Age at First Birth and Mother's Lifetime Earnings: Evidence from Danish Data. *PloS One*, *11*(1), e0146989.

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

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

McCrary, Justin. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, *142*(2), 698–714.

Neyt, Brecht, Vandenbulcke, Sarah, & Baert, Stijn. (2019). Are Men Intimidated by Highly Educated Women? Undercover on Tinder. *Economics of Education Review*.

Niederle, Muriel, & Vesterlund, Lise. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *Quarterly Journal of Economics*, *122*(3), 1067–1101.

Ong, David. (2016). Education and Income Attraction: An Online Dating Field Experiment. *Applied Economics*, *48*(19).

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

Ong, David, Yang, Yu, & Zhang, Junsen. (2015). Hard to Get: The Scarcity of Women and the Competition for High-income Men in Chinese Cities. *Working Paper*.

Parish, William L., Luo, Ye, Stolzenberg, Ross, Laumann, Edward O., Farrer, Gracia, & Pan, Suiming. (2007). Sexual Practices and Sexual Satisfaction: A Population Based Study of Chinese Urban Adults. *Archives of Sexual Behavior*, *36*(1), 5–20.

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

Pollet, Thomas V., & Nettle, Daniel. (2009). Partner Wealth Predicts Self-Reported Orgasm

Frequency in a Sample of Chinese Women. *Evolution and Human Behavior*, *30*(2), 146–151.

Qian, Yue, & Qian, Zhenchao. (2014). The Gender Divide in Urban China: Singlehood and Assortative Mating by Age and Education. *Demographic Research*, *31*, 1337–1364.

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

Ruffle, Bradley J., & Shtudiner, Ze'ev. (2014). Are Good-Looking People More Employable? *Management Science*, *61*(8), 1760–1776.

Shenhav, Na'ama. (2020). Lowering Standards to Wed? Spouse Quality, Marriage, and Labor Market Responses to the Gender Wage Gap. *Forthcoming Review of Economics and Statistics*.

Stockley, Paula, & Campbell, Anne. (2013). Female Competition and Aggression: Interdisciplinary Perspectives. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1631), 20130073.

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

Weiss, Yoram, Yi, Junjian, & Zhang, Junsen. (2019). *Hypergamy, Cross-boundary Marriages, and Family Behavior. International Economic Review*.

Wiswall, Matthew, & Zafar, Basit. (2018). Human Capital Investments and Expectations about Career and Family. *NBER Working Papers Series*.

Wright, Randall, Kircher, Philipp, Julien, Benoît, & Guerrieri, Veronica. (2020). Directed Search and Competitive Search: A Guided Tour. *Journal of Economic Literature, Forthcoming*.

Xu, Yujing, & Yang, Huanxing. (2016). Targeted Search with Horizontal Differentiation in the Marriage Market, (December), 1–53.

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

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

VIII. Tables and Figures

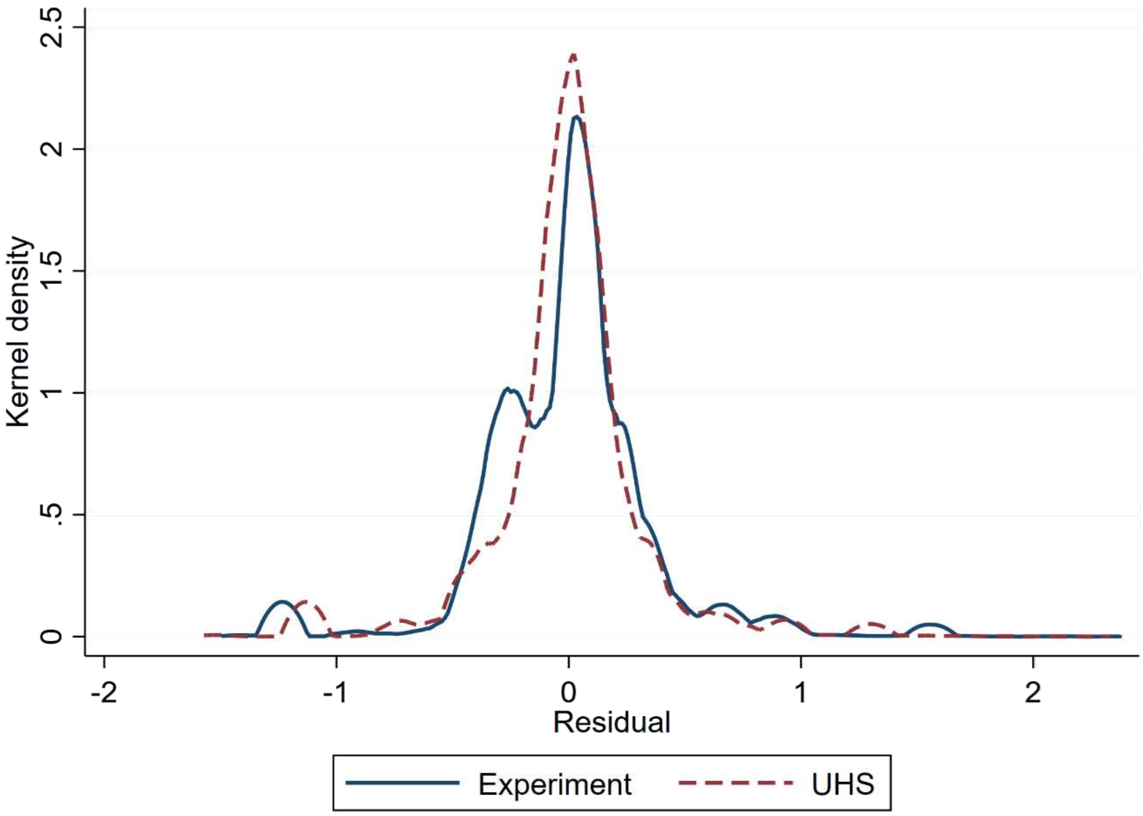


FIGURE 1: RESIDUAL DISTRIBUTION OF WOMEN'S VISITOR AND SINGLE UHS WOMEN'S INCOME

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

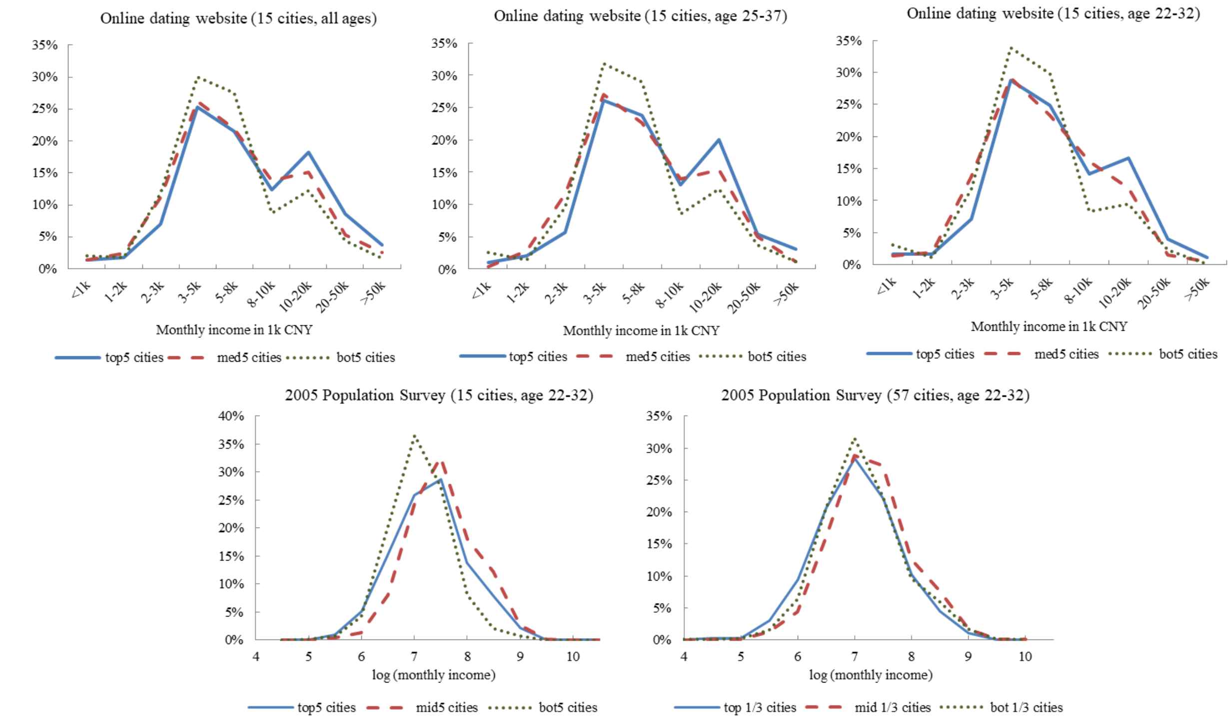


FIGURE 2: MEN'S INCOME DISTRIBUTIONS ON ONLINE DATING WEBSITE AND CENSUS DATA

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

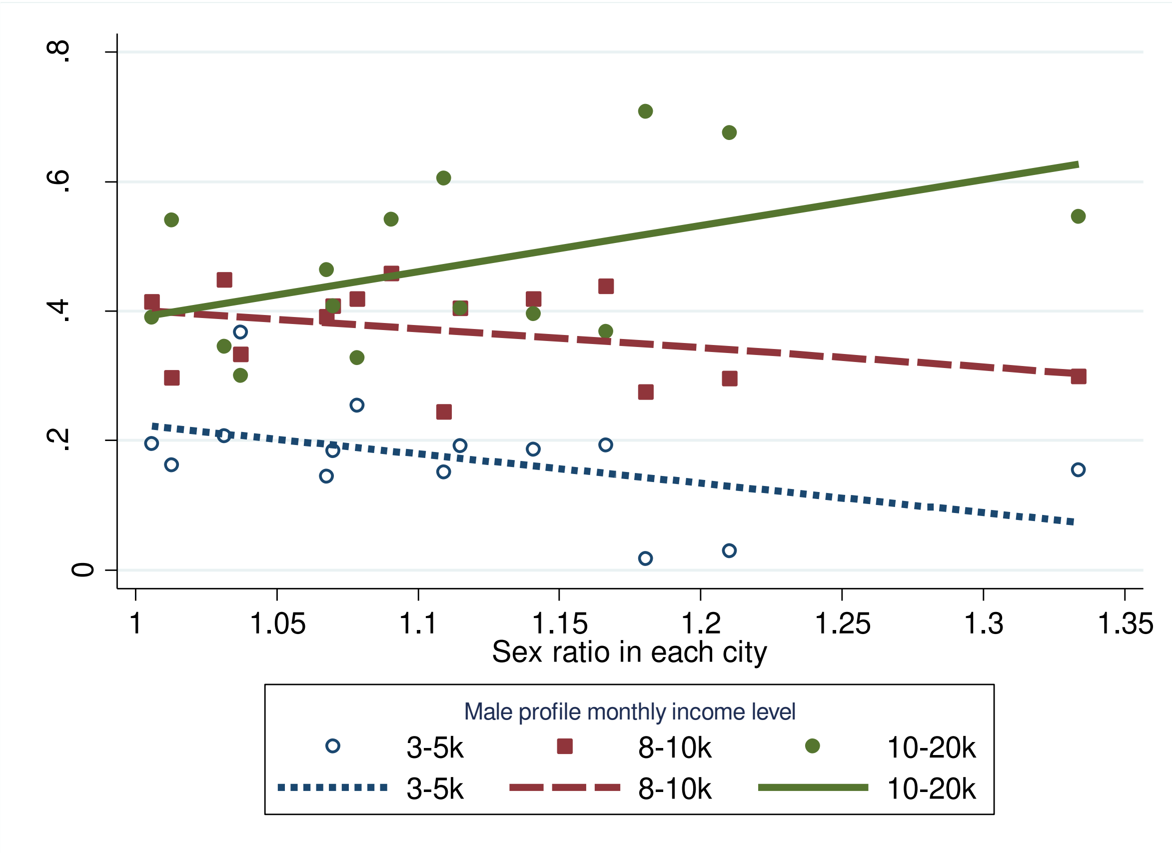


FIGURE 3: SHARE OF WOMEN'S VISITS BY MALE PROFILE INCOME LEVEL AND SEX RATIO

Notes: The women visitors to our male profiles between the age of 25-37 years old are between the ages of 18 and 45 years old. The horizontal axis indicates the sex ratio of the city. The vertical axis displays the share of visits, which is the total number of visits received by each income type of male profiles in each city divided by the visits to all our profiles for all male income types in the same city. The three lines represent the three income types of our male profiles. The graph shows an increasing gap between the visit rates of women for men who report an income 10-20*k* CNY/month and those who report 3-5*k* and 8-10*k* for men.

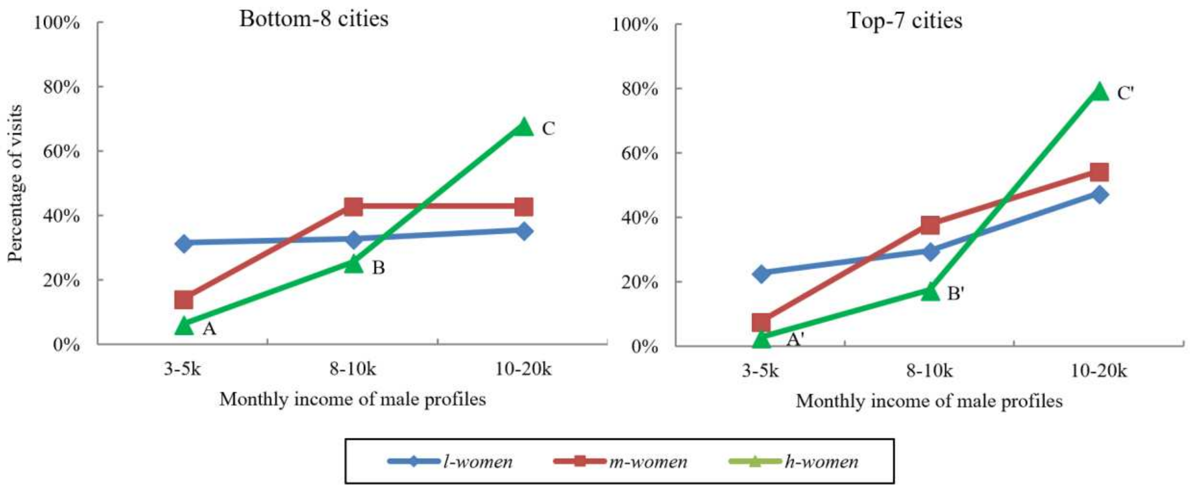


FIGURE 4: SHARE OF WOMEN'S VISITS TO MALE PROFILES BY WOMEN VISITOR'S INCOME AND SEX RATIO

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

TABLE 1: REGRESSION COMPARING WOMEN'S VISITOR CHARACTERISTICS AND UHS SAMPLE

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

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

TABLE 2: ORDERED PROBIT REGRESSION OF WOMEN'S VISITS ON MALE PROFILE INCOME

Dependent variable:	Income level of male profile visited (low (3-5k), middle (8-10k), high (10-20k))				
	Age 18-45	Age 18-45	Age 20-30	Age 20-30 (2SLS)	Age 31-45
	(1)	(2)	(3)	(4)	(5)
m-women dummy	0.156	-0.636	-0.993*	-0.889	-0.210
	(0.113)	(0.508)	(0.543)	(0.567)	(0.788)
h-women dummy	0.502***	-0.736*	-1.727***	-1.943***	0.117
	(0.159)	(0.393)	(0.620)	(0.714)	(0.754)
Sex ratio	**0.432**	**-6.219****	**-11.381*****	**-6.512**	**5.093**
	(0.638)	(2.491)	(3.485)	(5.336)	(3.577)
Sex ratio*m-women dummy	**0.891**	**5.401***	**10.285****	**8.854**	**-3.895**
	(0.859)	(3.222)	(5.157)	(5.469)	(6.485)
Sex ratio*h-women dummy	**1.765****	**10.879*****	**15.861*****	**15.826*****	**5.657**
	(0.892)	(2.678)	(4.770)	(5.773)	(5.285)
Beauty		-1.184	-1.973**	-1.142	2.006**
		(0.800)	(0.821)	(0.980)	(0.905)
Beauty*m-women dummy		1.172	1.713*	1.466	-0.133
		(0.967)	(1.008)	(1.042)	(1.634)
Beauty*h-women dummy		1.957***	3.722***	3.960***	-1.321
		(0.736)	(1.124)	(1.284)	(1.931)
Sex ratio*beauty		**12.131*****	**20.817*****	**13.027**	**-20.534*****
		(4.277)	(5.097)	(8.138)	(6.764)
Sex ratio*beauty*m-women dummy		**-8.339**	**-16.738***	**-13.786**	**14.794**
		(7.248)	(9.311)	(9.705)	(16.266)
Sex ratio*beauty*h-women dummy		**-19.219*****	**-29.824*****	**-28.938*****	**7.079**
		(4.912)	(7.906)	(9.855)	(12.547)
Additional controls:					
Age and education dummies of female visitors	Y	Y	Y	Y	Y
Mean and standard deviation of men's and women's incomes in each city		Y	Y	Y	Y
Observations	1,760	867	548	548	308
Pseudo R^2	0.049	0.066	0.085	0.085	0.088

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

TABLE 3: ORDERED PROBIT REGRESSION OF WOMEN'S VISITS ON MALE PROFILE INCOME

Dependent variable:	Income level of male profile visited (low (3-5k), middle (8-10k), high (10-20k))		
	Age 20-30	Age 20-30 (2SLS)	Age 31-45
	(1)	(2)	(3)
m-women dummy	-1.619	-1.027	-1.464
	(1.172)	(1.265)	(1.426)
h-women dummy	0.006	-0.308	-0.487
	(1.595)	(1.428)	(1.232)
Sex ratio	**-4.289**	**3.793**	**-3.215**
	(3.532)	(6.261)	(3.149)
Sex ratio*m-women dummy	**0.804**	**0.241**	**4.315*****
	(1.721)	(1.597)	(1.620)
Sex ratio*h-women dummy	**-1.962**	**-2.599**	**7.053****
	(1.953)	(1.665)	(2.954)
Beauty	-3.494*	-3.099	0.853
	(1.929)	(2.036)	(1.292)
Beauty*m-women dummy	4.001*	2.888	4.565*
	(2.263)	(2.359)	(2.603)
Beauty*h-women dummy	2.105	2.389	-0.497
	(2.913)	(2.653)	(2.237)
Sex ratio*beauty	**9.518***	**-3.466**	**-2.816**
	(5.680)	(10.222)	(6.693)
Mean income H-men	**-0.483****	**-0.649*****	**0.348**
	(0.201)	(0.208)	(0.237)
Mean income H-men*beauty	**0.567**	**0.868****	**-0.284**
	(0.402)	(0.403)	(0.495)
Mean income H-men*m-women dummy	**0.615*****	**0.674*****	**-0.137**
	(0.187)	(0.207)	(0.265)
Mean income H-men*beauty*m-women dummy	**-0.853****	**-1.001*****	**-0.426**
	(0.368)	(0.375)	(0.563)
Mean income H-men*h-women dummy	**0.990*****	**0.927*****	**-0.276**
	(0.288)	(0.279)	(0.260)
Mean income H-men*beauty*h-women dummy	**-1.370****	**-1.322*****	**0.084**
	(0.538)	(0.499)	(0.657)
Additional controls:			
Age and education dummies of female visitors	Y	Y	Y
Mean and standard deviation of men's and women's incomes in each city	Y	Y	Y
Observations	548	548	308
Pseudo R^2	0.098	0.098	0.117

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

TABLE 4: OLS REGRESSION OF MEN'S VISITS ON FEMALE PROFILE'S BEAUTY

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

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

TABLE 5: LOGIT REGRESSION OF WOMEN'S MARRIAGE PROBABILITY (CENSUS DATA)

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

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

TABLE 6: LOGIT REGRESSION OF WOMEN'S PROBABILITY OF MARRIAGE BY BEAUTY (CFPS DATA)

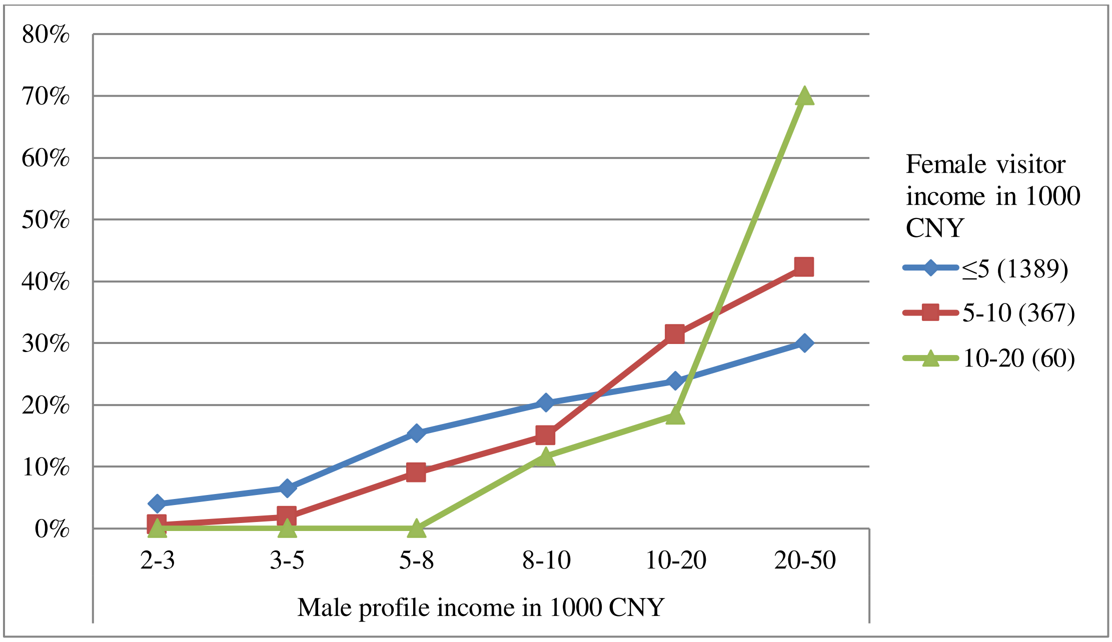
Dependent variable	
	Logit: 1 = married, 0 = single
Sex ratio	0.945
	(3.173)
h-women dummy	-0.332
	(0.305)
Sex ratio*h-women dummy	**-7.294****
	(3.183)
Beautiful dummy	0.823**
	(0.370)
Beautiful dummy**h*-women dummy	-0.607
	(0.423)
Sex ratio*beautiful dummy	-1.238
	(3.756)
Sex ratio*h-women dummy*beautiful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

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

Online Appendices

Appendix 1. Visit rates of men and women to online dating profiles with randomly assigned income (Reproduced from Ong and Wang (2015))

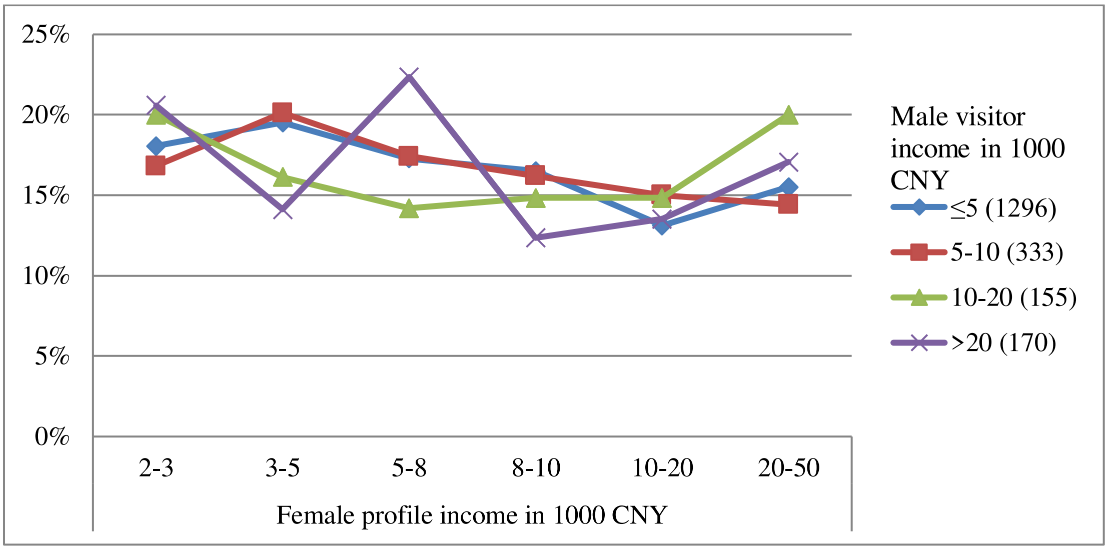
A-Figure 1 shows the share of clicks or "visits" by women of different income levels to male profiles with randomly assigned incomes. All three graphs show an increasing trend, which suggests that women not only visit high-income men with greater frequency, but specifically those who are higher income than themselves. The shallowest slope is for the women who report earning less than 5,000 CNY per month. The steepest slope is for the women who report earning 10,000-20,000 CNY per month. The kink in each graph that marks the statistically significant incremental increase in women's visit rates when the profile's income exceeds the average income of the women visitors. This kink further suggests women's ALM, along with the increase in the slope.



A-FIGURE 1: PERCENT OF FEMALE VISIT PER INCOME LEVEL VS. INCOME OF MALE PROFILES.

Notes: The number in brackets is the count of visits for women at the adjacent income level.

In contrast to the pattern in A-Figure 2, men of all income levels visit female profiles of all income levels with roughly equal probability.



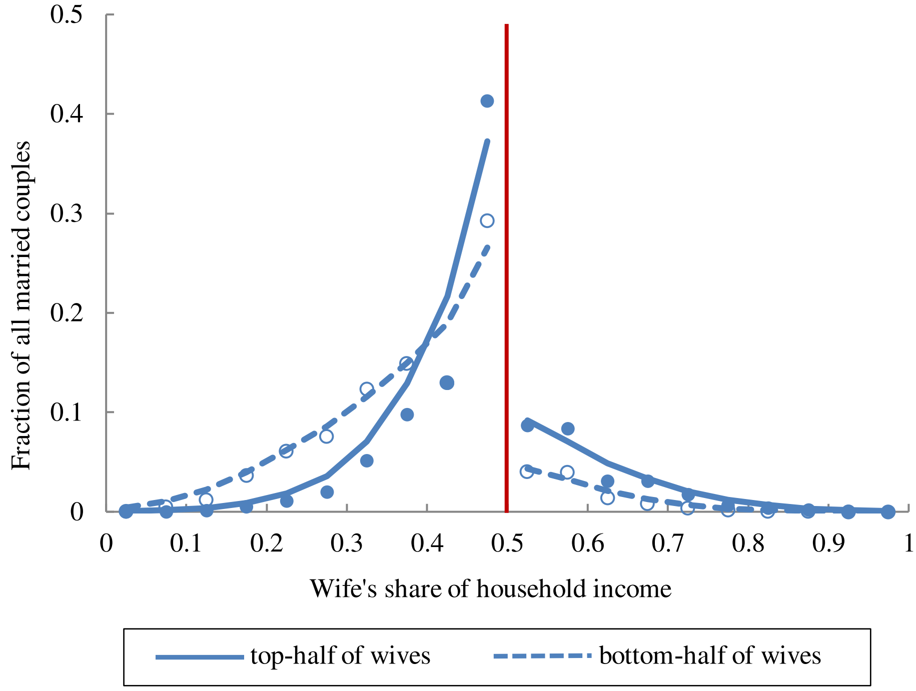
A-FIGURE 2: PERCENT OF MALE VISITS PER INCOME LEVEL VS. INCOME OF FEMALE PROFILES.

Notes: The number in brackets is the count of visits for men at the adjacent income level.

Appendix 2. Marriage shares (*Reproduced from Ong, Yang, and Zhang* (2015))

We find further support for the ALM hypothesis suggested by the online dating data in Appendix 1 with the 2005 1 Percent Population Survey (often called "the mini-Census") data of married couples. We restrict the sample to people between 22 and 55 years old with urban *hukou* and then identify and match 39,988 first-marriage couples in nuclear families where both spouses have positive incomes. A-Figure 3 depicts the distribution of married couples by the wife's relative income (wife's income divided by the sum of the husband's and wife's incomes), following the approach of Bertrand et al. (2015), using U.S. data. We extend their results by ranking the wife's income within each city and age, and divide them into the top- and the bottom-half groups, represented by the solid and dashed lines in A-Figure 3, respectively, to separate possibly different effects across high- and low-income women.

In each graph, we confirm with Chinese data the extreme "skewness" in the distribution of marriages, as measured by a sharp drop-off in the frequency of marriages just as the wife's income exceeds that of her husband's. The McCrary (2008) test of the discontinuity of the distribution function indicates that both distributions drop at the 0.5 point of household income ($p < 0.01$). Our findings are similar to those of Bertrand et al. (2015). They interpret the drop-off in their data as evidence of an aversion among couples for the wife earning more than the husband. However, they did not categorize the wives into high- and low-income groups and compare these drops.



A-FIGURE 3: DISTRIBUTION OF SPOUSES BY WIFE'S SHARE OF HOUSEHOLD INCOME WITH CITY LEVEL DATA

Notes: Data are from China 2005 1 Percent Population Survey. We identified and matched 39,988 first-marriage couples in nuclear families, both between 22 and 55 years old with urban *hukou* and positive incomes. Wife's share of household (monthly) income = wife income / household income, grouped into 20 bins. We estimated the discontinuity to the *right* of the point of 0.5. Solid and dash lines represent top-and bottom-half of wives by income within each age in each city.

The gap between the dashed line (bottom-half of wives by income) and the solid line (top-half of wives by income) in A-Figure 3, when we disaggregate by income, is revealing. We briefly discuss here how the increasing drop supports our women's ALM hypothesis. See Ong, Yang, and Zhang (2015) for further details.

Starting from the left-hand side of the 0.5 share of household income point, the solid line is noticeably below the dashed line and then crosses to rise above it. This crossing indicates that, before the 0.5 share point, a smaller fraction of high-income wives (solid line) contribute a small share to household income and a larger fraction of high-income wives contribute more equally with their husbands to their household income. The probability drop at 50 percent of the household income point is 0.253 for the bottom-half of women, increasing to 0.326 for the top-half women. The Kolmogorov–Smirnov test confirms that the difference in the overall distributions of the two groups is significant ($p < 0.01$). The general trend of the drop in the distribution as the wife's income approaches the 50 percent point of the household income is even more evident in Appendix 7 A-

Figure 3 of Ong, Yang, and Zhang (2015), in which we plot the size of the drop against the wife's percentile income rank.

The larger drop in the marriage rate as the women's income exceed that of her husband for high-income women is consistent with high-income women being even more willing to settle for men with only slightly higher or equal income than themselves than low-income women. We expect such a pattern, given our ALM hypothesis, because high-income women face competition for high-income men from low-income women, while low-income women do not face similar competition from high-income women for low-income men. This higher level of competition constrains high-income women to a lower average match surplus from their mate's income (based on their reservation value as singles) than low-income women. Under our ALM hypothesis, this reservation value is correlated with the women's own income. This lower surplus is reflected in a larger mass concentrated on the immediate left of the 0.5 share point of household income, and therefore, a larger drop off to the right of 0.5 for the high-income women group as compared to the low-income women group. These trends are confirmed in A-Figure 4, where we divide the women into income percentiles.

A-FIGURE 4: SIZE OF DROP IN DISTRIBUTION OF SPOUSES BY WIFE'S INCOME RANK WITH CITY LEVEL DATA

Notes: The percentile rank of wife's (monthly) income is calculated for each city and age. For all wives within a percentile rank, we plot their relative income distribution, like A-Figure 3, and estimate the distribution drop at the 0.5 point. Then we plot the distribution drops vs. the income percentile ranks.

Appendix 3. A Game Theoretic Illustration of the Competition between High- and Low-income Women[1]

The model in A-Table 1 illustrates how beautiful-looking low-income women adjusting their tradeoff between a sure match with a low-income man and an uncertain match with a high-income man can crowd out plain-looking high-income women, who (because of their ALM) seek these high-income men exclusively, from the marriage market. We model only this across-income competition between beautiful low-income women and plain-looking high-income women, and not the within-income competition, because our main goal here is to show that high-income women can be negatively affected by the increase in the sex ratio or income of high-income men due to the attendant increase in the entry of beautiful low-income women. We focus on the latter women also because we showed empirically that plain-looking low-income women decrease their search intensity for high-income men, and also that beautiful high-income women's probability of marriage did not decrease with the increase in sex ratio. Including these two types of women in our model would not add insight into the crowding out of the plain-looking high-income women (hereafter, high-income women) by the beautiful low-income women (hereafter, low-income women) and may obscure the insights on the crowding out effect.

We adapt the directed search model of homogenous buyers and sellers in Burdett et al. (2001) to the case of heterogenous types in the marriage market. (See Section III Conceptual Framework for a discussion of our motivation for using the directed search framework.) Burdett et al. focus on equilibrium prices, which determine the surplus share between homogenous buyers and sellers. In our marriage-matching context, we can think of the women as buyers and the men as sellers. Instead of paying a price, women pay a search cost which varies with their income type. Making the buyer's cost exogenous and not equivalent to the seller's benefit does not substantially change the analysis from Burdett et al., because in our setting, the men (i.e., the sellers) are passive players. However, the men's preference for beauty affects the probability of a woman matching with them. Burdett et al. also assume symmetric numbers of buyers and sellers. We discuss how changes in the sex ratio would affect our results in our simple setting. See

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

also the 2 x 2 model on p. 27 in Wright et al.'s (2020) review. Eeckhout and Kircher's (2010) generalize much of the directed search literature, including Burdett et al. (2001), to continuous actions and any distribution of a continuum of heterogeneous types (Chade et al., 2017; Wright et al., 2020). [2]

To specify the payoffs in this game, recall that Becker (1973) posits a complementarity between the wife's beauty and the husband's income which is supported empirically (Weiss et al., 2019).[3] Thus, one possible formulation of the matching incentives is as follows. The choice of a woman of income type $w \in \{l, h\}$ who earns y_w, where l represents low-income, and h represents high-income, and has beauty percentile rank $0 < b_w < 1$ to be single or to marry a man of income type $m \in \{L, H\}$ who earns x_m, where L represents low-income, and H represents high-income, is

$$u_w = max\{y_w,\ x_m b_w + \delta(x_m + f(n, b_w) y_w)\}\,. \qquad \text{Eq. (1)}$$

The first component in the max function is the value of being single, which is just the woman's own income. The second component in the max function is the value of a possible marriage for the woman. The product $x_m b_w$ reflects the complementarity discussed above and may include the public good of children in the family. δ, $0 \leq \delta \leq 1$, is the woman's share of the household income as wife, which we assume, for simplicity, is exogenously given by the social norms, i.e., is independent of the woman's income or beauty. Women discount their own after-marriage wages by $f(n, b_w)$, $0 \leq f(n, b_w) \leq 1$. The discount factor, $f(n, b_w)$, declines in b_w and the man's or the woman's gender identity n.[4] Thus, a woman may marry a man who earns x_m if the value of marriage is higher than being single:

[2] Lang, Manove, and Dickens (2005) model racial discrimination within a directed search framework can be extended to model our marriage market by incorporating four instead of two types of searchers. However, in contrast to our marital context, their model would still not include heterogenous search costs from marriage market frictions, heterogenous outside options from different opportunity costs of marriage (which drives our ALM), and complementarity between the searcher's and their target's characteristics (i.e., female beauty and male income).

[3] See Section 4 of Part I of the Appendix of Becker (1973). This complementarity between men's income and women's beauty may be due to the complementarity between male resource investment and female fertility for viable offsprings. In terms of the attractiveness of high resource males to females, beyond the economics literature cited in the introduction indicating that women prefer higher income men (Hitsch et al., 2010a; Ong & Wang, 2015), there is an extensive evolutionary psychology literature on how conspicuous consumption on the part of men may increase their attractiveness to women by signaling the men's income and potential investment of resources in their offspring. In terms of the attractiveness of indicators of female fertility to males, facial femininity, which adds to female facial beauty and attractiveness to men, signals high levels of the female hormone oestrogen, and therefore, fertility (Jokela, 2009; Rhodes, 2006). There is evidence that men's fertility increases on their income (Hopcroft, 2006).

It is obvious that men find beautiful women sexually attractive. As to evidence that women find high-income men sexually attractive, the female orgasm is hypothesized to be a potential discriminator of male quality in terms of the likelihood of conception or selective bonding with higher quality sires (Pollet & Nettle, 2009). Female orgasm frequency is an important component of female sexual satisfaction and it increases on the husband's income based on the analysis of a large representative sample from the China Health and Family Life Survey (Parish et al., 2007; Pollet & Nettle, 2009).

[4] This decline reflects past findings suggesting that men and women tend to discount women's after-marriage income (Hitsch et al., 2010a; Ong & Wang, 2015), especially after childbirth (Wiswall & Zafar, 2018) and our hypothesis that this is due to the labor

$$x_m b_w + \delta(x_m + f(n, b_w)y_w) \geq y_w \, . \qquad \text{Eq.(2)}$$

Rearranging, we have the following necessary condition

$$x_m \geq \frac{1-\delta f(n,b_w)}{b_w+\delta} y_w \qquad \text{Eq.(3)}$$

for marriage. ALM means that a woman would reject a man if his income is less than hers. A sufficient condition to prevent $x_m \leq y_w$ when Equation (3) holds is to require the coefficient $\frac{1-\delta f(n,b_w)}{b_w+\delta}$ for the pair to be larger than 1.[5]

We incorporate this utility into a simple model based on a standard directed search framework and show how low-income women switching from surer matches with low-income men to uncertain matches with high-income men, when the sex ratio and the value of marrying such high-income men increases, may crowd out high-income women from the marriage market. We model the competitive search between high- and low-income women for high-income men as the competition between two types of players: a high-income woman (Column Player) and a low-income woman (Row Player) (see the game matrix in A-Table 1). We model the search choices of each player for 2 x 2 outcomes in this game in terms of whether to exert the extra effort cost necessary to potentially match with the high-income man: $\{(Try, Not\ Try) \ \times \ (Effort, No\ effort)\}$.[6] The payoffs for the low-income woman are the first coordinates of each pair for each outcome as represented in the above matrix, whereas that of the high-income woman is the second.

---

market opportunity costs of women's childbearing. Gender identity can exacerbate the decline either alone (Bertrand et al., 2015) or interacted with women's fertility (Jeon & Ong, 2018).

5 ALM means that a woman would reject a man if his income is less than hers. Here, we formulate the man's incentive to reject a woman as a potential wife to show why men may not have an ALM for their wife's income. The incentive of the man of income type $m \in \{L, H\}$ who earns x_m, where L represents low-income, and H represents high-income, to be single or to marry a woman who earns y_w and has beauty b_w would be

$$u_m = max\{x_m, \ x_m b_w + (1-\delta)(x_m + f(n, b_w)y_w)\} \, . \qquad \text{Eq.}(f1)$$

We assume, as in women's surplus function, a parallel complementarity between men's income and women's beauty/fertility expressed in $x_m b_w$ in men's surplus from marriage. This assumption is supported by our online dating finding that men's preference for beautiful women increases with their income, i.e., they anticipate a higher surplus from more beautiful women. Similar to the choice of women, men may accept a woman who earns y_w and has beauty b_w if

$$x_m b_w + (1-\delta)(x_m + f(n, b_w)y_w) \geq x_m \qquad \text{Eq.}(f2)$$

For the empirical finding of a lack of ALM on the part of men to hold, it is sufficient that the beauty of the potential wife must compensate for her potential share of the husband's income ($b_w > \delta$). In that case, he may accept her regardless of her income y_w. Otherwise, the wife's income must compensate for the difference between the share of her share of his income δ and her beauty rank b_w, discounted by his share of her after-marriage income $(1-\delta)f(n, b_w)$:

$$y_w \geq \frac{\delta - b_w}{(1-\delta)f(n,b_w)} x_m . \qquad \text{Eq.}(f3)$$

Equation $f3$ suggests a potential explanation for why men tend to ignore women's reported or potential income in online dating studies (Neyt, Vandenbulcke, & Baert, 2019; Ong, 2016). Men and women tend to visit profiles that not only have a chance of making an acceptable match, but are also even more desirable than their own (Bruch & Newman, 2018). Therefore, if men focus on profiles with high facial beauty, then b_w would tend to be greater than δ.

6 Eeckhout and Kircher (2010) incorporate such entry costs into the equilibrium tradeoffs between the value and the probability of matching when they relate their work to prior literature.

To simplify our analysis of the equilibrium payoffs and avoid unnecessarily complicated algebra, we denote the payoff of a woman of income type $w \in \{l, h\}$ marrying a man of income type $m \in \{L, H\}$ as $\theta_w(m)$, where $\theta_w(m) = x_m b_w + \delta(x_m + f(n, b_w) y_w)$, based on Equation (3). Importantly, $\theta_w(m)$ increases on x_m; both types of women get greater surplus matching with the high-income man.

A-TABLE 1: GAME MATRIX FOR COMPETITION BETWEEN HIGH- AND LOW-INCOME WOMEN FOR HIGH-INCOME MEN

		High-income woman (Plain-looking)	
		$No\ effort\ (1-e)$	$Effort\ (e)$
Low-income woman (Beautiful-looking)	$Try\ (t)$	$\theta_l(H) - c_l, y_h$	$z\theta_l(H) + (1-z)\theta_l(L) - c_l$, $zy_h + (1-z)\theta_h(H) - c_h$
	$Not\ Try\ (1-t)$	$\theta_l(L), \theta_h(H)$	$\theta_l(L), \theta_h(H) - c_h$

The low-income woman automatically "gets" the low-income man[7] or she can Try for the high-income man who earns x_H. She values marriage to him at $\theta_l(H)$ minus her cost of effort (from searching or otherwise, e.g., putting more effort in grooming) $c_l \geq 0$. The high-income woman values the high-income man at $\theta_h(H)$ and can put in $Effort$ at cost $c_h \geq 0$ in getting him. We can write $\theta_h(H) = a \cdot \theta_l(H)$, where a is a positive constant.[8] Hence, we need only specify variations in $\theta_l(H)$ when we do comparative statics. To capture the greater frictions faced by the low-income woman in meeting a high-income man due to sorting by coeducational and market institutions, we assume that: $c_l > c_h \geq 0$.

Starting from the outcome $(Not\ Try, No\ Effort)$ in the lower left hand corner of A-Table 1, if neither woman wants to compete for the high-income man, we assume that the high-income woman gets him. This last assumption adds to the lower friction ($c_l > c_h$) that the high-income woman enjoys in getting the high-income man, e.g., she is in the same coeducational institution or the same firm. Going to the diagonal outcomes $(Try, No\ Effort)$ and $(Not\ Try, Effort)$, if one of the two women is competing to get

7 Recall that an estimated 30 million women are missing from the prime age marriage market. Low-income men are especially desperate given women's preference for high-income men. Hence, women can presumably match with low-income men at low cost.

8 In any case, the precise value does not affect our results.

the high-income man and the other is not, the one who competes succeeds for sure, and the other gets the payoff of their next best option, as described next.

On the top right of A-Table 1 is the outcome $(Try, Effort)$ where both women compete for the high-income man. We capture the relative attractiveness of the low-income woman as compared to the high-income woman to the high-income man with the variable z. z represents the odds that the low-income woman gets the high-income man when she chooses Try and the high-income woman also chooses $Effort$. Though none of our qualitative results depend on the value of z, the greater beauty of the low-income woman implies $z > \frac{1}{2}$. Each woman gets her next best option when she fails to get the high-income man. For the low-income woman, this is the low-income man, who she then gets with probability $(1 - z)$. For the high-income woman, this is her singlehood payoff y_h, which she gets with probability z.

To simplify the payoffs, we normalize high-income woman's income $y_h = 0$ from this point forward. We further assume $\theta_w(H) > \theta_w(L)$ for both $w \in \{l, h\}$, that is, getting the high-income man yields a higher payoff than getting the low-income man. Also, assume $\theta_l(H) - c_l > 0,\ z\theta_l(H) + (1 - z)\theta_l(L) - c_l > 0$ and $z(1 - z)\theta_h(H) - c_h > 0$ so that the payoff in each case is non-negative.

The two pure strategy equilibria are $(Not\,Try, No\,Effort)$, which occurs for low values of $\theta_l(H)$, and $(Try, Effort)$, which occurs for high values of $\theta_l(H)$. Relating these equilibria to the low-income woman crowding out the high-income woman, in the equilibrium in which the low-income woman chooses Try, the value of the high-income man, $\theta_l(H)$, is high enough for the low-income woman to switch from a sure match with the low-income man (who she values at $\theta_l(L)$) to an uncertain match with the high-income man (who she values at $z\theta_l(H) - c_l$). This switch forces the high-income woman to compete for the high-income man and choose $Effort$. Her probability of getting him decreases from 1 to $1 - z$, resulting in her being crowded out with probability z. The parameter values necessary and sufficient for these equilibria are summarized in A-Table 2.

Next, we look for the interior mixed strategy equilibrium. Let e stand for the probability high-income woman chooses $Effort$ and t stand for the probability that the low-income woman chooses Try. 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

$$(1-e)(\theta_l(H)-c_l)+\ e(z\theta_l(H)+(1-z)\theta_l(L)-c_l)=\theta_l(L) \qquad \text{Eq.(4)}$$

$$t(z\cdot 0+(1-z)\theta_h(H)-c_h)+(1-t)(\theta_h(H)-c_h)=t\cdot 0+(1-t)\theta_h(H) \qquad \text{Eq.(5)}$$

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

In A-Table 2, if $\theta_l(H)$ is below $\theta_l(L)+c_l$, then it is a dominant strategy for the low-income woman to $Not\ try$. The high-income woman's payoff is $\theta_h(H)$. When the value of trying for the high-income man for the low-income woman, $z\theta_l(H)+(1-z)\theta_l(L)-\ c_l$, rises above $\theta_l(L)$, the low-income woman chooses Try with strictly positive probability. In doing so, she switches her search target from the low-income man (who she can always get) to an uncertain match with the high-income man. This 'competitive entry' of the low-income woman forces the high-income woman to compete for the high-income man by choosing $Effort$. The probability with which the low-income woman gets the high-income man rises from zero to z. Her payoff is still $\theta_l(L)$ since she is still indifferent in this mixed strategy equilibrium to her old payoff. The probability with which the high-income woman's gets the high-income man decreases from 1 to $1-z$. Her payoff drops discontinuously from $\theta_h(H)$ to $\theta_h(H)-\frac{c_h}{1-z}$. If $\theta_l(H)$ increases further to above $\theta_l(L)+\frac{c_l}{z}$, the low-income woman tries with probability 1 (full entry), and high-income woman's payoff again drops discontinuously, this time from $\theta_h(H)-\frac{c_h}{1-z}$ to $(1-z)\theta_h(H)-c_h$.These equilibrium strategies and payoffs are detailed in A-Table 2 and are further illustrated in A-Figure 5.

We need only change the interpretation of this game slightly to model the effect of an increase in sex ratio in the directed search framework, where agents target their search based on expected utility. Let the two types of women now be two populations of otherwise homogenous individual women, namely high- and low-income.[9] We interpret the probability distribution of their equilibrium strategies as the share of each type of women adopting these strategies. We also interpret z, which previously represents the probability of one woman getting one high-income man, to now represent the *share* of high-income men that the low-income women population gets, given the shares of both the high- and the low-income women populations that put in $Effort$ or Try, respectively. The share of low-income women matching is $z \cdot s$, where s represents the sex ratio. The share of high-income women matching is $(1-z) \cdot s$. We can relabel $z_l(s) = z \cdot s$ and $z_h(s) = (1-z) \cdot s$. When sex ratio increases from s to s', the probability which enters into each type of woman's expected value of searching for the high-income men would be $z_l(s')$, $z_l(s') > z_l(s)$, and $z_l(s')$, $z_h(s') > z_h(s)$, rather than $z_l(s)$ and $z_h(s)$, respectively. Hence, given $z_l(s)$ and $z_h(s)$, the expected value of searching for high-income men is

$$z_l(s) \cdot \theta_l(H) = (z \cdot s) \cdot \theta_l(H) = z \cdot (s \cdot \theta_l(H)) \qquad \text{Eq.(6)}$$

for low-income women, and

$$z_h(s) \cdot \theta_h(H) = ((1-z) \cdot s) \cdot \theta_h(H) = (1-z) \cdot (s \cdot \theta_h(H)) \qquad \text{Eq.(7)}$$

for high-income women. When $z_l(s)$ and $z_h(s)$ increase with the sex ratio, the expected value of searching for high-income men also increases, just as if the value of matching with these men increases to $s \cdot \theta_l(H)$ and $s \cdot \theta_h(H)$ while the share of the high-income men each income type of women gets is fixed at z and $1-z$.

It is obvious that we would find similar results with an increase in s that we find with an increase in $\theta_l(H)$ and $\theta_h(H)$. Hence, the effect of an increase in the sex ratio is similar to an increase in the income of high-income men because the expected value which determines the player's decision includes both the value and probability of potential matches.

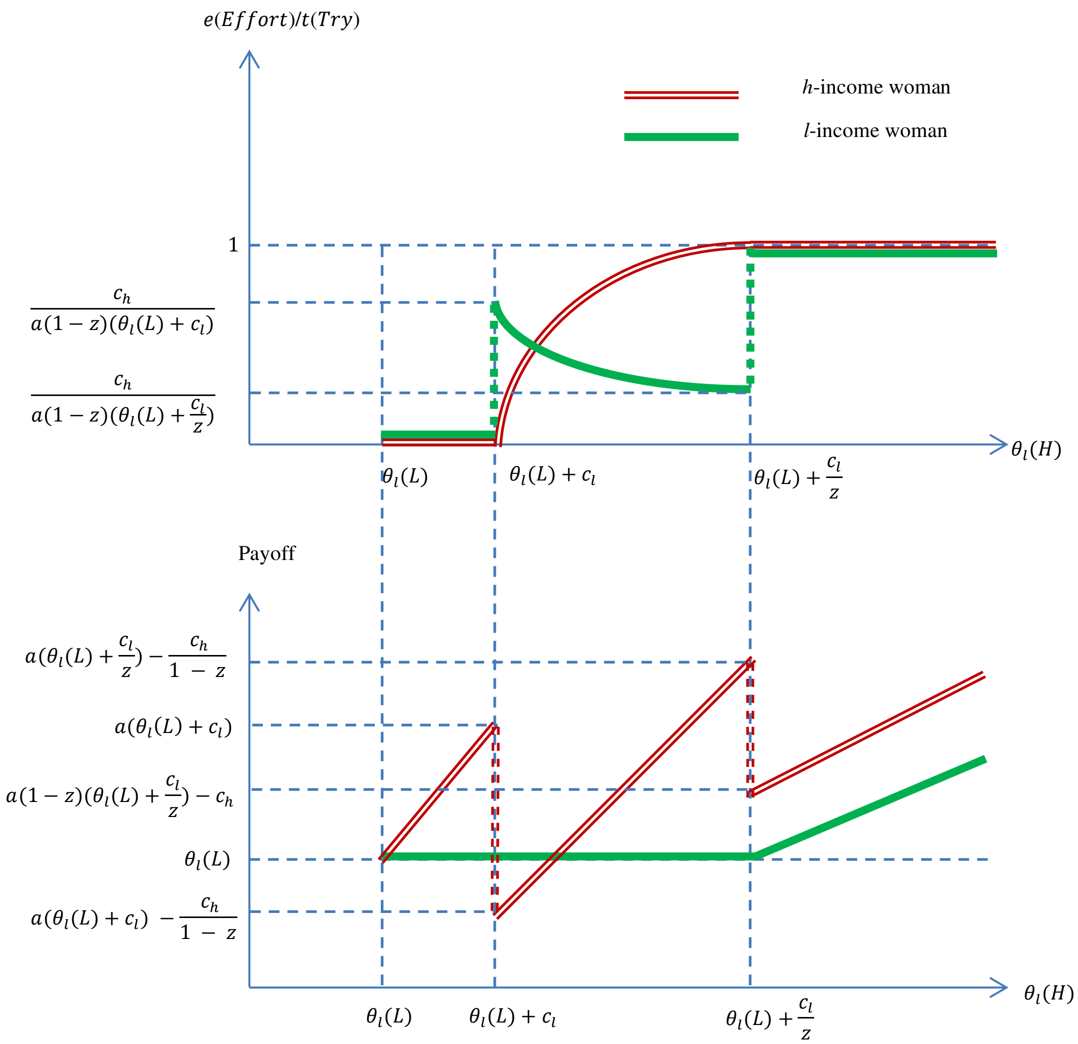
---

[9] The probability with which individual players choose an action in a mixed strategy Nash equilibrium can be interpreted as shares of a population of players choosing pure strategies. See the purification theorem in Harsanyi (1973) for details.

A-TABLE 2: EQUILIBRIUM PAYOFFS FOR EACH TYPE OF WOMAN GIVEN z AND $\theta_l(H)$

	$\theta_l(L) < \theta_l(H) < \theta_l(L) + c_l$	$\theta_l(L) + c_l < \theta_l(H) < \theta_l(L) + \frac{c_l}{z}$	$\theta_l(H) > \theta_l(L) + \frac{c_l}{z}$
h-woman			
$e(Effort)$	0	$\frac{\theta_l(H) - \theta_l(L) - c_l}{(1-z)(\theta_l(H) - \theta_l(L))}$	1
payoff	$\theta_h(H)$	$\theta_h(H) - \frac{c_h}{1-z}$	$(1 - z)\theta_h(H) - c_h$
l-woman			
$t(Try)$	0	$\frac{c_h}{(1\ -\ z)\cdot\theta_h(H)}$	1
payoff	$\theta_l(L)$	$\theta_l(L)$	$z\theta_l(H) + (1 - z)\theta_l(L) - c_l$

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



A-FIGURE 5: STRATEGIES AND PAYOFFS FOR HIGH- AND LOW-INCOME WOMEN

Notes: We impose $\theta_h(L) = a\theta_l(L)$ for simplicity. z is the probability that the l-type woman gets the high-income man when both types of women search for him. The top panel illustrates the equilibrium strategy for the high (h)- and low (l)-income women, whereas the bottom panel illustrates their respective payoffs. The thin double-lines are those of the h-income woman. The thick green lines are those of the l-income woman. The low-income woman switches from her sure payoff with matching with the low-income man to her uncertain payoff by trying to match with the high-income man when $z\theta_l(H) + (1-z)\theta_l(L) - c_l > \theta_l(L)$ forcing the h-woman to compete for the high-income man.

We now conduct a numerical simulation to illustrate that our results from the two income and beauty types case hold still for a continuum of income and beauty types.

We simulate 1000 men and 1000 women. Men's income x_m follows a uniform distribution over [1.5, 2.5]. Women's income y_w follows a uniform distribution over [1, 2], where we allow for a gender wage gap. Women's beauty percentile b_w is assumed to be independent of her income and we randomly assign b_w from a uniform distribution over [0, 1]. When marrying a man with income x_m, the woman's marital surplus is given by

$$u_w(x_m) = 0.25 + 0.5x_m b_w + 0.5(x_m + 0.5y_w) \quad \text{Eq.(8)}$$

which is assumed to be deterministic for simplicity. A woman's value of being single is just her own income y_w. When married with a woman with (y_w, b_w), a man's marital surplus is given by

$$u_m(y_w, b_w) = 0.25 + 0.5x_m b_w + 0.5(x_m + 0.5y_w) + \varepsilon_m \quad \text{Eq.(9)}$$

where ε_m is an i.i.d. shock following the Type-I extreme value (TIEV) distribution. When there is a queue of women searching for a man, this TIEV shock term in men's surplus gives us the man's probability of choosing a particular woman (y_w, b_w) from the queue, which takes the convenient logit form:

$$p_m(y_w, b_w) = \frac{exp(u_m(y_w, b_w))}{\sum_w exp(u_m(y'_w, b'_w))} \quad \text{Eq.(10)}$$

where the denominator is the sum of the exponential of the man's potential surplus when matching with each woman in the queue. Note that the probability of a woman being chosen from the queue equals her match probability with man m, p_w, thus

$$p_w(x_m) = p_m(y_w, b_w) \quad \text{Eq.(11)}$$

We assume the search cost of women decreases with the difference between the woman's own income rank and the income rank of the man she searches for. We use income rank to capture the search costs from labor market segmentation. The search cost c_w is given by

$$c_w(x_m) = 0.5\frac{|rank_w(y_w) - rank_m(x_m)|}{1000-1} \quad \text{Eq.(12)}$$

where $rank_m(\cdot)$ and $rank_w(\cdot)$ are the income rank among men and women, respectively. The lowest income has a rank of 1, and the highest income has a rank of 1000. The function yields a search cost range of [0, 0.5]. The choice of coefficient 0.5 is to make the search cost non-trivial such that women would not be able to search without limit across income groups. On the other hand, the search cost should not be too high to prevent any entry of low-income women into the market for high-income men even when these men are much richer or more plentiful.

Given these elements, a woman's expected value from searching for a man who earns x_m is

$$p_w(x_m)u_w(x_m) - c_w(x_m) \qquad \text{Eq.(13)}$$

In equilibrium, a woman chooses to search for a man to maximize her expected value, taking as given all other women's searches. If the highest expected value from search is lower than the value of being single, she will not search and remain single. We simulate each woman's search choice and iterate until convergence to the pure strategy equilibrium in this discrete choice search model.[10]

The simulation starts with a default search: each woman searches for the man with the same income rank as hers at zero search cost. Thus, in this initial search, income is perfectly positively assortative. This is to reflect the educational and labor market segmentations in real life. Taking the queue at each man in the default search as given, we solve the first woman's maximization problem to decide which man she should search for. After the first woman adjusts her search, the second woman takes the updated queues as given and adjusts her search accordingly. We iterate this simulation until no woman has a profitable deviation from her last round's search target and we reach the equilibrium.

We simulate the equilibrium search targets and matching probabilities of different women and examine how they change with sex ratio or men's income distribution. In the baseline setting, we simulate 1000 men and 1000 women (i.e., sex ratio = 1) and let men's income follow a uniform distribution over [1.5, 2.5].

We first examine the effect of an increase in the sex ratio. We conduct three simulations, the baseline with sex ratio set at 1, another with sex ratio increased to 1.1 and third with sex ratio set at 1.2, all three keeping men's income distribution fixed. A-Figure 6

[10] Simulating mixed equilibrium strategies would be much more difficult since it requires solving and tracking the presumably continuous choice vector of search allocations of each woman's search probability across all men. To keep the simulation of equilibrium trackable, we leave that for future work.

compares women's search target in equilibrium between these simulations and the baseline. In A-Figure 6, we report the average income of the men searched by women (vertical axis) by their own income (horizontal axis) and beauty (bottom 1/3 in panel A, middle 1/3 in panel B, top 1/3 in panel C). Note that both women's income and beauty are drawn from uniform distributions, but we divide women into three equal subgroups by beauty for easier presentation. The dotted line in each panel represents the baseline equilibrium where sex ratio is set at 1. As shown in the figure, women's search is positively assortative on income in all three panels for this baseline case. The dashed and solid lines represent the equilibrium when sex ratio is 1.1 and 1.2, respectively.

In the baseline (dotted lines), no woman, regardless of her beauty, deviates from the default assignment where each woman searches for the man with same income rank as herself. No woman switches in the baseline case because given our parametrization of the search cost, switching to searching for a higher-income man and competing with the woman who initially searches for that man would incur both a higher search cost and a lower match probability that would not be compensated for by the higher value of the new man.

We next compare the search behavior of women when the sex ratio increases from 1 (dotted baseline) to 1.1 (dashed line). Starting with panel C, which graphs the search behavior of beautiful women for whom the effect of sex ratio is most pronounced, we see that the left end of the dashed line shifts upward compared to the dotted baseline. This upward shift by the left end indicates that the average income of the men being searched for by these low-income women increases. This upward shift in the targeted men's income is largest for the lowest-income women, because their initial matches have the lowest income. Thus, the opportunity cost for these low-income women (who give up a match with low-income men) in switching to searching for high-income men is the lowest.

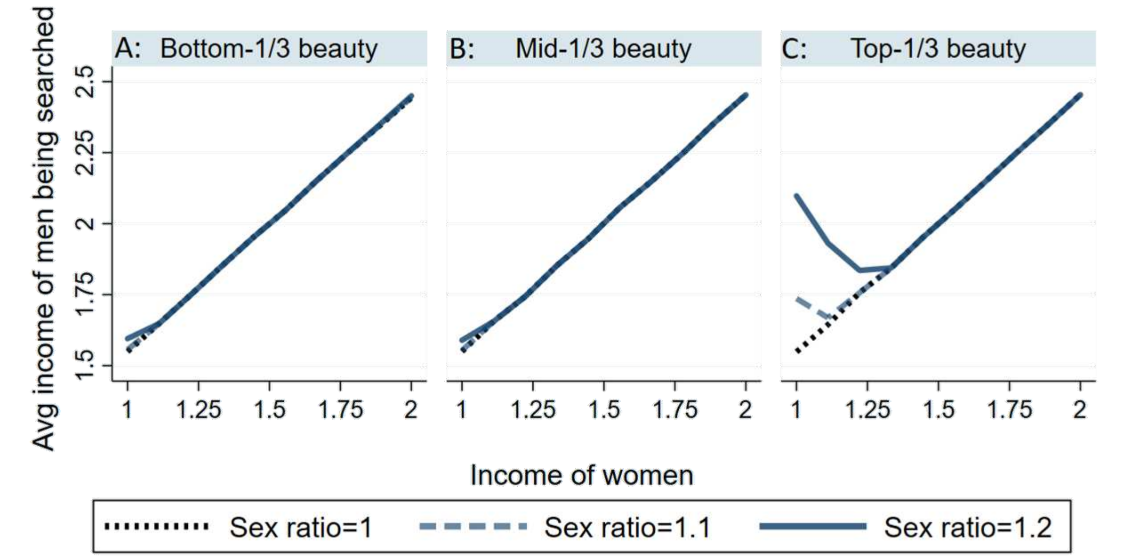
When the sex ratio increases further to 1.2 (solid line), the left end of the solid line in panel C shifts upwards even more significantly than the dashed line when the sex ratio was 1.1. Due to the complementarity between women's beauty and men's income in their marital surplus, beautiful low-income women most readily relinquish their initial match with a low-income man in the baseline case and switch their search target to high-income men.

For higher income women (e.g., those whose income is above 1.25), the income of their initial matches in the baseline case (their opportunity cost) is also higher. Thus, competing for higher-income men yields a smaller gain in terms of men's income and may entail a large decrease in match probability. For many of these higher income women, the switch to searching for even higher income men would generally result in a loss compared to staying with their initial sure match. Consequently, fewer of these women switch their searches to high-income men, and thus, the average income of men being searched (the vertical axis) is lower than that of the women with the lowest income levels.

Graphically, this effect of higher opportunity costs of switching is shown in the magnitude of the gap between the dashed line and the dotted baseline decreasing and eventually disappearing as women's income increases beyond 1.25.[11] Moreover, we see that despite the competitive entry of beautiful low-income women, high-income women also do not deviate from their searches for high-income men to searching for low-income men (i.e., the graphs are never decreasing on the women's income for high-income women), due to their aversion to marrying lower-income men.

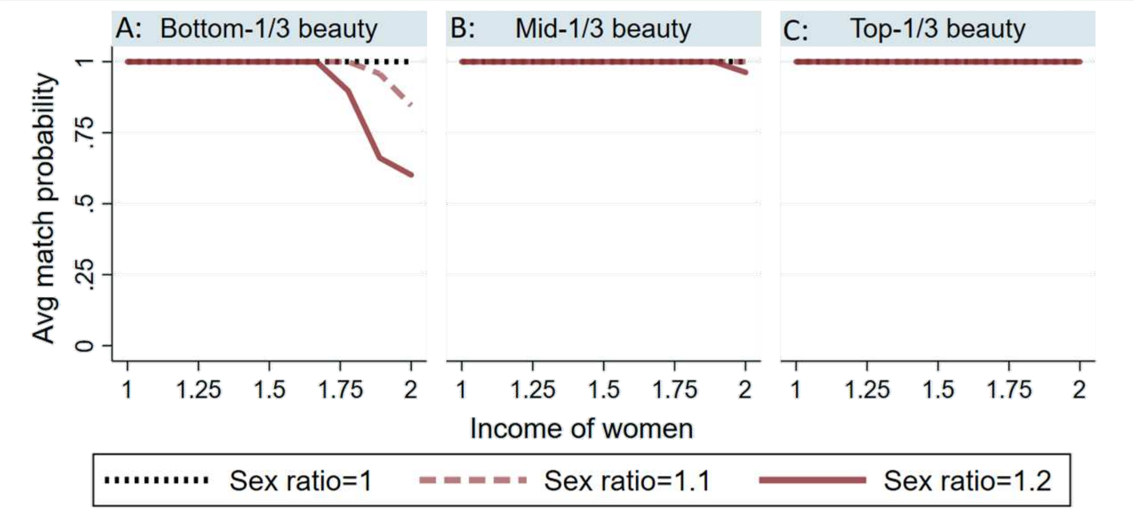
Moving right to left across panels in A-Figure 6, we see that in contrast to beautiful low-income women in panel C, low-income women of mid-1/3 (dashed line in panel B) and bottom-1/3 beauty (dashed line in panel A) increase their searches for high income men less significantly as sex ratio increases. This less pronounced increase is because of their lower gain in marital surplus from their lower beauty. The increase in the expected surplus from the increase in the sex ratio is not enough to compensate for the higher search cost and lower match probability with high-income men from the increased competition for these men. As a consequence, the dashed line overlaps with the dotted baseline in panels A and B for women of all income levels except for the very lowest near 1. This high degree of overlap is in contrast to panel C, where a divergence in search behavior is evident for women facing different sex ratios above the income level of 1.25. As in Panel C, high-income women in panels A and B continue to search for high-income men until the expected value from marital search becomes lower than the value of being single rather than switch to searching for low-income men.

[11] Note that we are less interested in the shape of any curve (e.g. the V-shape of the dashed and solid curves in Panel C), which is a comparison of women with different incomes cross-sectionally. Our focus is to examine the search strategy of any specific woman when the sex ratio increases, which is also the more appropriate way of interpreting the simulation results. Thus, essentially, we are looking at the change in search strategy from one curve to another curve vertically when the sex ratio increases, for any given woman.



A-FIGURE 6: SIMULATION OF SEX RATIO INCREASE: WOMEN'S EQUILIBRIUM SEARCH TARGET

Notes: This figure compares the equilibrium in three simulations. First, we simulate 1000 men and 1000 women in the baseline (sex ratio = 1, represented by the dotted lines). Second, we increase the sex ratio to 1.1 (the dashed lines). Third, we increase the sex ratio to 1.2 (the solid lines). All other settings are the same: men's income follows a uniform distribution over [1.5, 2.5]; women's income follows a uniform distribution over [1, 2] and their beauty follows a uniform distribution over [0, 1]. For easier presentation, we divide women into three equal subgroups by each third of beauty in each sub-figure.



A-FIGURE 7: SIMULATION OF SEX RATIO INCREASE: WOMEN'S EQUILIBRIUM MATCH PROBABILITY

Notes: This figure compares the equilibrium in three simulations. First, we simulate 1000 men and 1000 women in the baseline (sex ratio = 1, represented by the dotted lines). Second, we increase the sex ratio to 1.1 (the dashed lines). Third, we increase the sex ratio to 1.2 (the solid lines). All other settings are the same: men's income follows a uniform distribution over [1.5, 2.5]; women's income follows a uniform distribution over [1, 2] and their beauty follows a uniform distribution over [0, 1]. For easier presentation, we divide women into three equal subgroups by each third of beauty in each sub-figure.

A-Figure 7 compares women's marriage probability between the baseline sex ratio and two simulations with higher sex ratios. It reports women's average match probability in equilibrium (vertical axis) by their own income (horizontal axis) and beauty (in panels A, B and C). In the baseline case (dotted lines) when sex ratio is set at 1, given our parametrization, no woman has a sufficient incentive to deviate from her initial match. Thus, there is no competition among the women, and all the women match with probability 1 in all three panels.

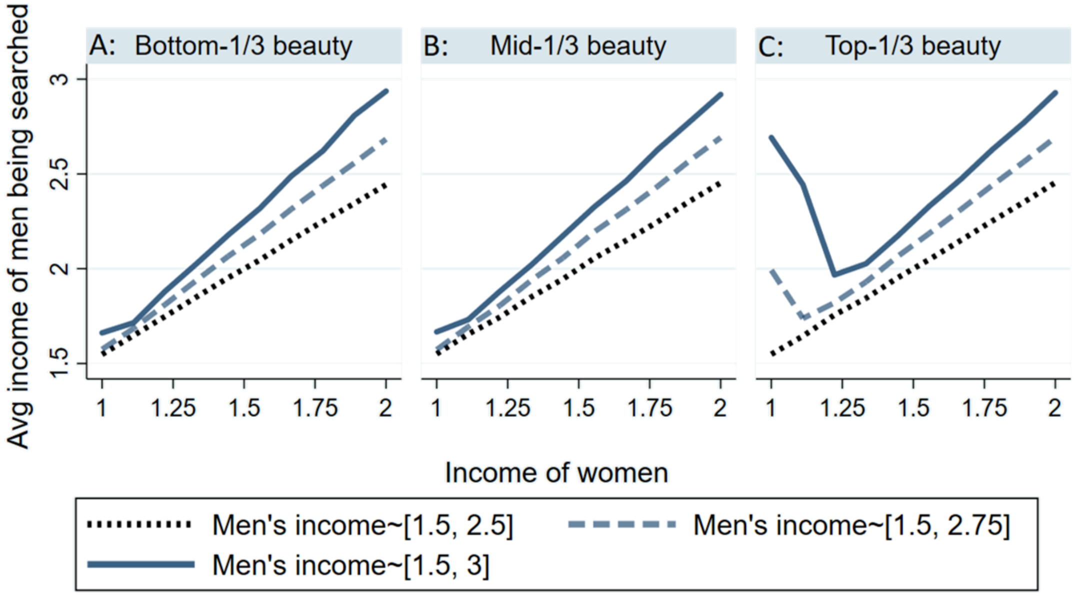
When the sex ratio increases to 1.1, the separation between the dashed lines and the dotted line in panels A and B shows the effect of beautiful low-income women's entry into the high-income men submarket. There is no separation in panel C. The decrease in match probability of high-income women is most pronounced for plain-looking high-income women because of the complementary between men's income and women's beauty in men's marital surplus and because of women's ALM. As a consequence of the complementarity of male income and female beauty in their surplus function, men will choose beautiful women with higher probability than plain-looking women when given a choice. Plain-looking high-income women, unlike plain-looking low-income women, do not substitute towards low-income men due to their binding ALM. They would rather resort to their singlehood outside option.

As a result of both of the complementarity and ALM effects, the match probability of plain-looking high-income women (right end of the dashed line in panel A) drops compared to the baseline, while that of the medium-beauty high-income women (right end of the dashed line in panel B) and beautiful high-income women (right end of the dashed line in panel C) do not drop. For low-income women, whose ALM is not binding, their match probability is constant across all three beauty groups.

When the sex ratio increases to 1.2 (solid lines), beautiful low-income women's competitive entry is stronger. Consequently, the match probability of plain high-income women drops more substantially (right end of the solid line in panel A), which means more plain high-income women are crowded out from the marriage market.

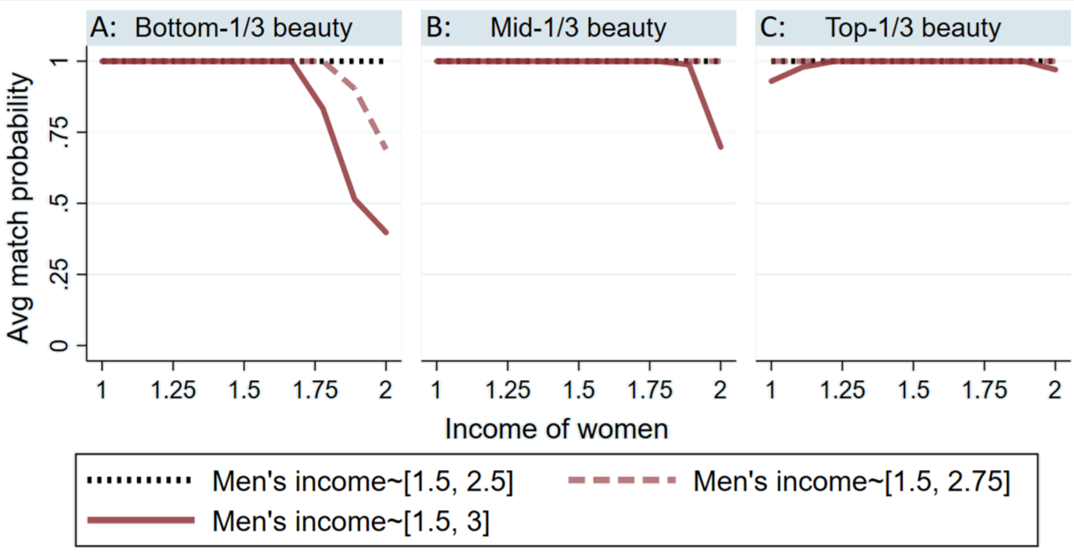
Next, we fix the sex ratio to 1 and conduct two additional simulations in which we increase men's income. Compared to the baseline, the first additional simulation increases the upper bound of men's income distribution from 2.5 to 2.75 and the second one further increases it to 3. We compare them with the baseline equilibrium in A-Figure 8 and A-Figure 9.

A-Figure 8 reports the average income of men being searched in equilibrium and A-Figure 9 reports women's match probability. The dotted line represents the baseline, and the dashed and solid lines represent the simulation with men's income distribution over [1.5, 2.75] and [1.5, 3], respectively. Again, women are divided into three subgroups by their beauty.



A-FIGURE 8: SIMULATION OF MEN'S INCOME INCREASE: WOMEN'S EQUILIBRIUM SEARCH TARGET

Notes: This figure compares the equilibrium in three simulations. First, we simulate 1000 men and 1000 women in the baseline where men's income follows a uniform distribution of [1.5, 2.5] (represented by the dotted lines). Second, we increase men's income distribution to [1.5, 2.75] (the dashed lines). Third, we increase men's income distribution to [1.5, 3] (the solid lines). All other settings are the same: sex ratio is 1; women's income follows a uniform distribution of [1, 2] and their beauty follows a uniform distribution of [0, 1]. For easier presentation, we divide women into three equal subgroups by each third of beauty in each sub-figure.



A-FIGURE 9: SIMULATION OF MEN'S INCOME INCREASE: WOMEN'S EQUILIBRIUM MATCH PROBABILITY

Notes: This figure compares the equilibrium in three simulations. First, we simulate 1000 men and 1000 women in the baseline where men's income follows a uniform distribution of [1.5, 2.5] (represented by the dotted lines). Second, we increase men's income distribution to [1.5, 2.75] (the dashed lines). Third, we increase men's income distribution to [1.5, 3] (the solid lines). All other settings are the same: sex ratio is 1; women's income follows a uniform distribution of [1, 2] and their beauty follows a uniform distribution of [0, 1]. For easier presentation, we divide women into three equal subgroups by each third of beauty in each sub-figure.

In A-Figure 8, when men's income distribution is increased to [1.5, 2.75], the dashed lines in three panels all shift upwards, reflecting the fact that all women's initial search targets become richer under the new income distribution. As with increases in the sex ratio, while plain (dashed line in panel A) and medium-beauty (dashed line in panel B) women's searches are still positively assortative on income, beautiful low-income women (left end of the dashed line in panel C) give up their initial match of low-income men and switch their search target to richer men. Again, this is because, in the context of the complementarity between male income and female beauty, an increase in men's income leads to the largest increase in surplus for the most beautiful women, so they are most willing to bear the higher search cost and lower match probability (compared to their original sure match) to compete for these men.

Similar to the effect of an increase in the sex ratio, the upward shift of the search target's income is largest for the lowest-income women in panel C, due to their lowest opportunity cost of giving up initial matches. When men's income distribution is further increased to [1.5, 3] (solid lines), men at the top of the distribution are even more valuable. Beautiful low-income women's entry (left end of the solid line in panel C) is stronger, as expected, compared to the dashed line, and even some medium-beauty low-income women (left end of the solid line in panel B) switch to search for richer men.

Moving to A-Figure 9, as a result of beautiful low-income women's strong entry into the submarket for rich men when men's income distribution increases from [1.5, 2.5] to [1.5, 2.75], the match probability of plain high-income women (right end of the dashed line in panel A) drops considerably. In contrast, the match probability of beautiful low-income women (left end of the dashed line in panel C) remains close to 1 after the entry, because rich men reciprocate more to beautiful women (due to the complementarity of male income and female beauty in the marital surplus). When men's income distribution increases further to [1.5, 3] (solid lines), plain and medium-beauty high-income women are crowded out even more from the marriage market. This additional crowding-out is reflected in the sharper drop in the match probability in the right region of panels A and B.

To sum up, the simulation with continuous income types for men and income and beauty types for women illustrates that, in the context of the directed search, when high-income men are more plentiful or richer, plain high-income women may be crowded out from the marriage market by the competitive entry of beautiful low-income women. This result is consistent with the result obtained from the 2 x 2 discrete type example discussed earlier.

Appendix 4. Background on Cities and Visitors

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

A-TABLE 3: CHARACTERISTICS OF CITIES USED IN THE ONLINE DATING EXPERIMENT

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

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

A-TABLE 4: REGRESSION OF MEN'S MEAN INCOME ON LOCAL SEX RATIO WITH CITY-LEVEL DATA

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

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

A-TABLE 5: REGRESSION OF MEN'S INCOME DISPERSION ON LOCAL SEX RATIO WITH CITY LEVEL DATA

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

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

A-TABLE 6: SUMMARY STATISTICS OF AGE, INCOME, AND EDUCATION FOR MALE VISITORS

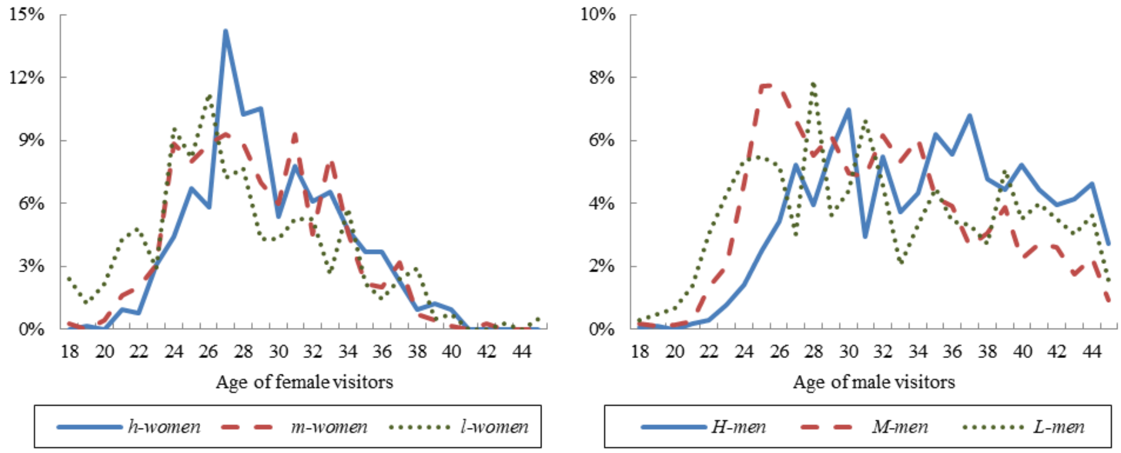
Male	Obs.	Mean	Std. Dev.	Min	Max
Age	5981	33.93	7.580	18	69
Income (1*k* CNY)	5706	10.39	11.03	1	50
Education (years)	5705	15.14	1.689	12	21

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

A-TABLE 7: SUMMARY STATISTICS OF AGE, INCOME, AND EDUCATION FOR FEMALE VISITORS

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

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



A-FIGURE 10: AGE DISTRIBUTION OF WOMEN'S VISITS TO MALE PROFILES AND MEN'S VISITS TO FEMALE PROFILES

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

Appendix 5. Instrumental Variable Robustness Check

A-TABLE 8: THE FIRST STAGE REGRESSION FOR IV-ORDERED PROBIT REGRESSION IN TABLES 2 AND 3

Dependent variable:	Sex ratio (log)	
	For column (4) of Table 2	For column (2) of Table 3
	(1)	(2)
m-women dummy	-0.086*	0.028
	(0.050)	(0.029)
h-women dummy	-0.055	0.049
	(0.062)	(0.036)
Bartik sex ratio	**0.917*** **	**1.263*** **
	(0.100)	(0.063)
Bartik sex ratio**m*-women dummy	0.277**	-0.056
	(0.115)	(0.047)
Bartik sex ratio**h*-women dummy	0.214	-0.069
	(0.138)	(0.050)
Beauty ranking	-0.191**	-0.020
	(0.075)	(0.049)
Beauty ranking**m*-women dummy	0.144	-0.045
	(0.092)	(0.053)
Beauty ranking**h*-women dummy	0.069	-0.050
	(0.111)	(0.060)
Bartik sex ratio*beauty ranking	0.492***	0.008
	(0.182)	(0.105)
Bartik sex ratio*beauty ranking**m*-women dummy	-0.475**	
	(0.216)	
Bartik sex ratio*beauty ranking**h*-women dummy	-0.312	
	(0.248)	
Mean income of *H*-men		0.022***
		(0.004)
Mean income of *H*-men*beauty ranking		-0.002
		(0.007)
Mean income of *H*-men**m*-women dummy		0.006
		(0.004)
Mean income of *H*-men*beauty ranking**m*-women dummy		-0.013
		(0.008)
Mean income of *H*-men**h*-women dummy		0.006
		(0.006)
Mean income of *H*-men*beauty ranking**h*-women dummy		-0.006
		(0.010)
Additional controls:		
Age and education dummies of female visitors	Y	Y
Mean and standard deviation of men's and women's incomes in each city	Y	Y
Constant	-0.055	-0.058*
	(0.042)	(0.030)
Observations	548	548
F-statistic	115.4	167.3
R^2	0.866	0.936

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