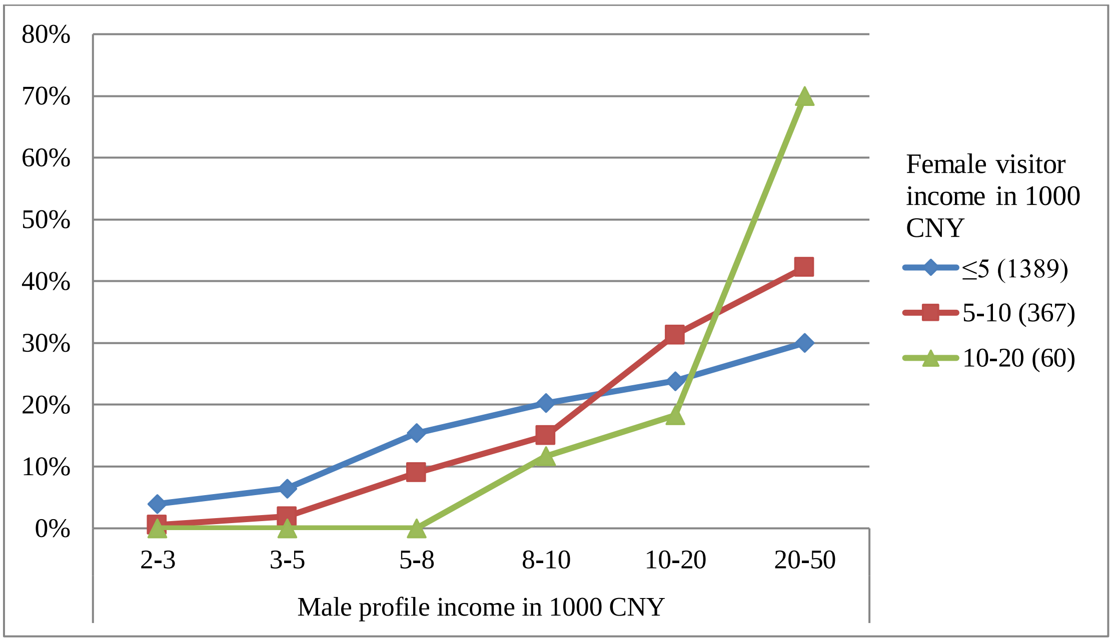
Online Appendices

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

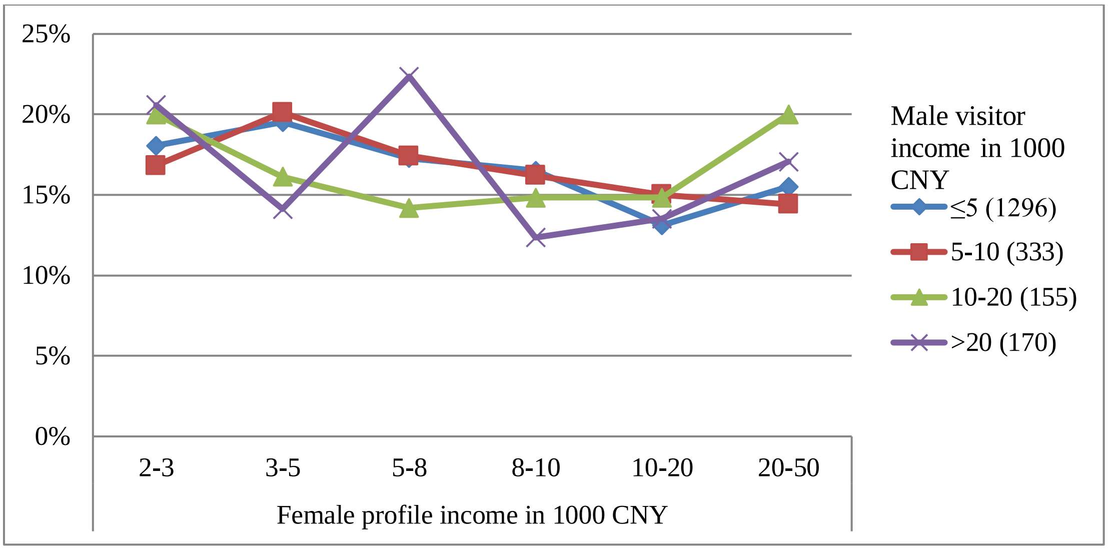
A-Figure 1 shows the share of clicks or "visits" by women of different income levels to male profiles with randomly assigned incomes. All three graphs show an increasing trend, which suggests that women not only visit high-income men with greater frequency, but specifically those who are higher income 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.



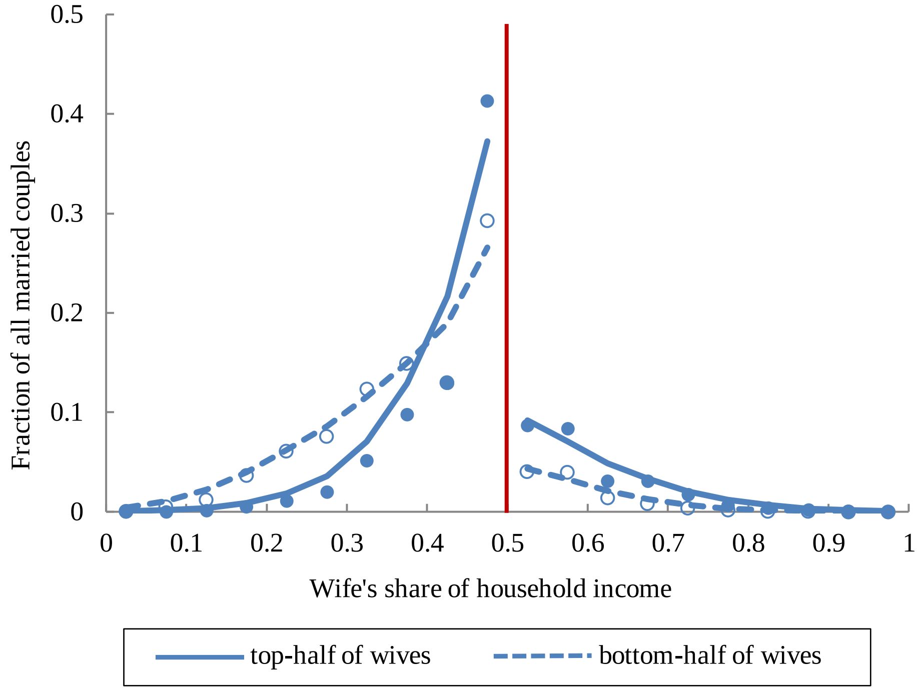
A-FIGURE 2: PERCENT OF MALE VISITS PER INCOME LEVEL VS. INCOME OF FEMALE PROFILES.

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 (2019) 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, Eeckhout, & Smith, 2017; Wright et al., 2019). [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, Yi, & Zhang, 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, Hortaçsu, & Ariely, 2010; 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., 2010; 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 $1 - z$.

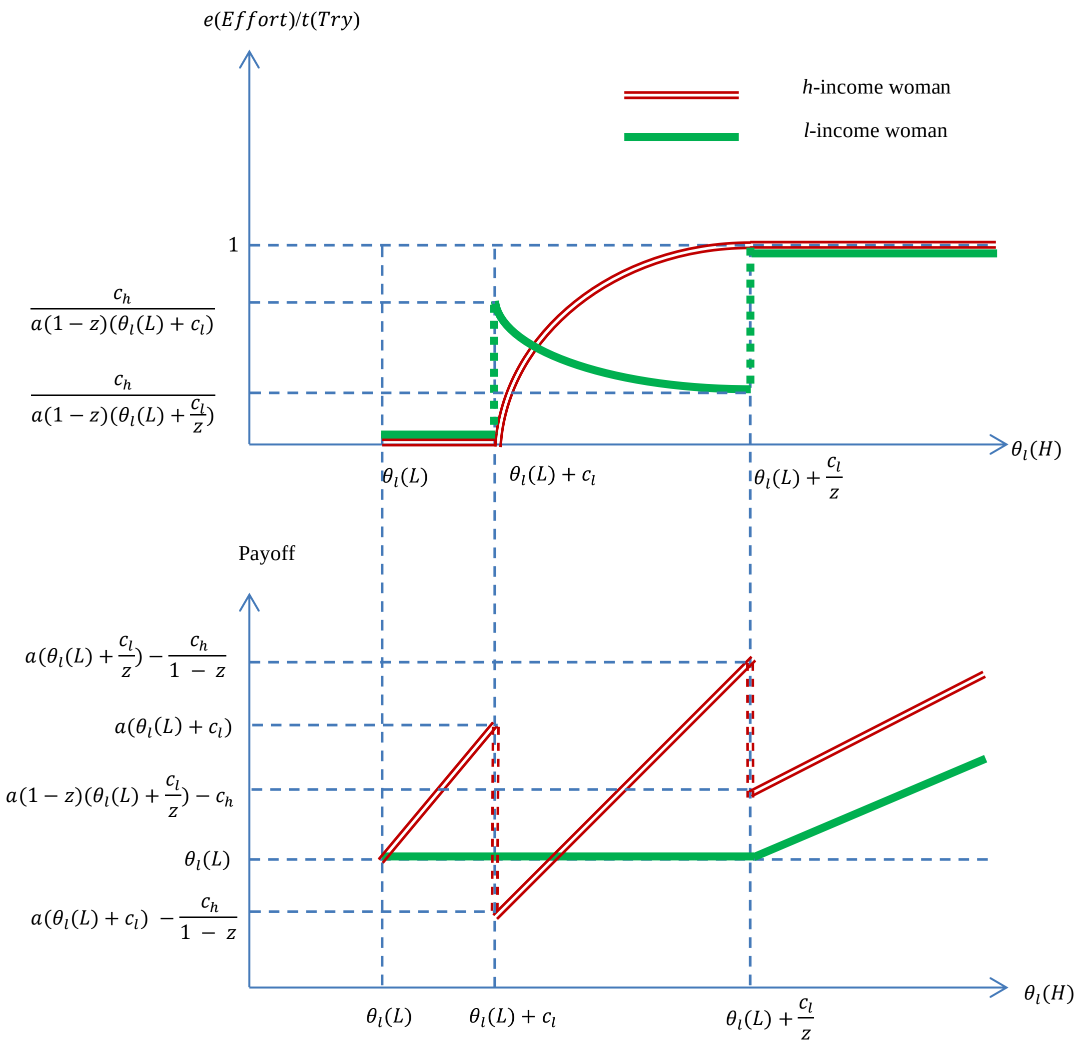
It is obvious that we would find similar results with an increase in s that we find with an increase in $\theta_l(H)$ and $\theta_h(H)$. Hence, the effect of an increase in the sex ratio is similar to an increase in the income of high-income men because the expected value which determines the player's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

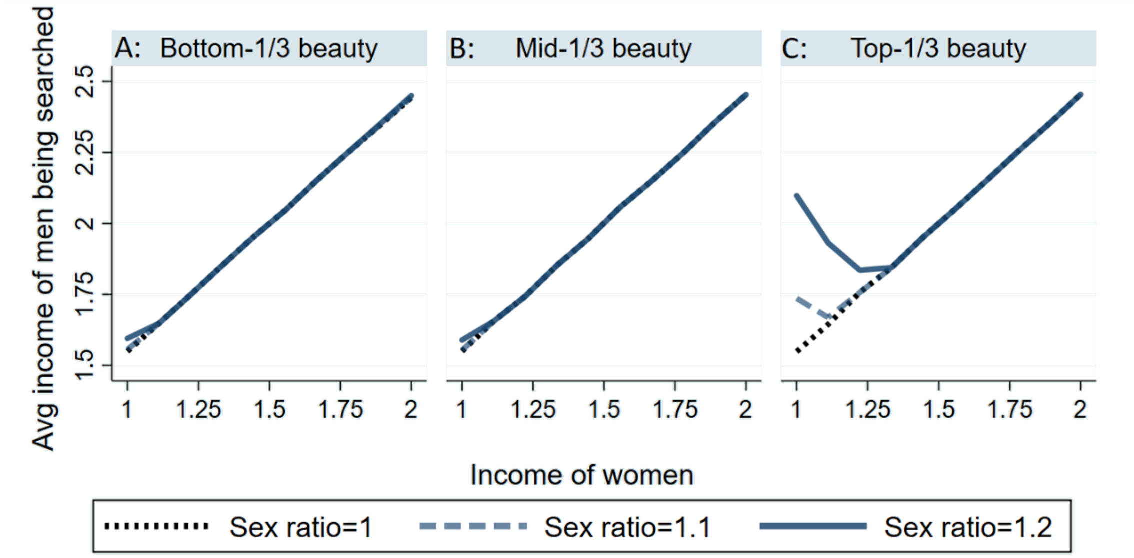
When the sex ratio increases further to 1.2 (solid line), the left end of the solid line in panel C shifts upwards even more significantly than the dashed line when the sex ratio was 1.1. Due to the complementarity between women's beauty and men's income in their marital surplus, beautiful low-income 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.

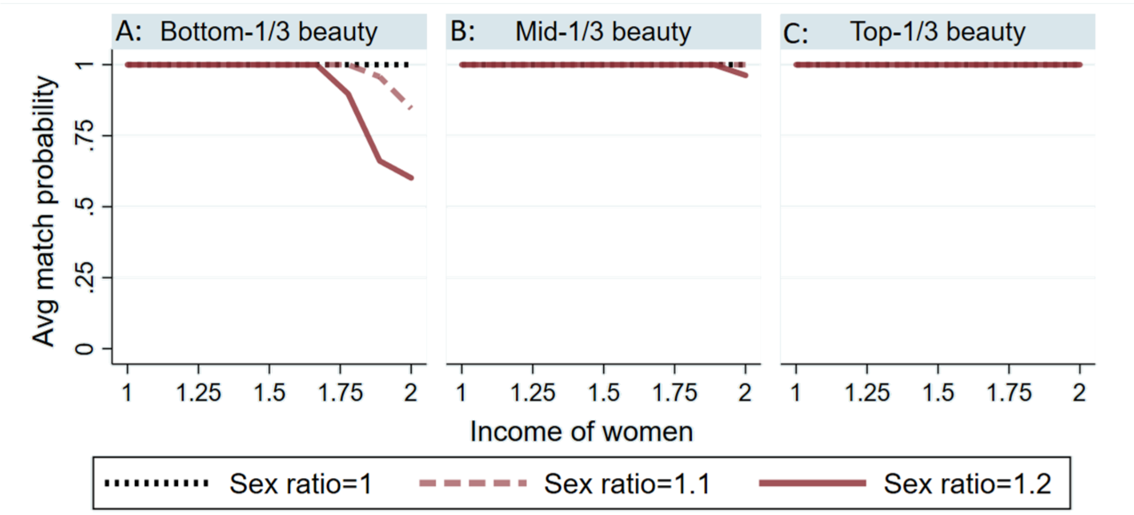
Moving right to left across panels in A-Figure 6, we see that in contrast to beautiful low-income women in panel C, low-income women of mid-1/3 (dashed line in panel B) and bottom-1/3 beauty (dashed line in panel A) increase their searches for high income men less significantly as sex ratio increases. 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.

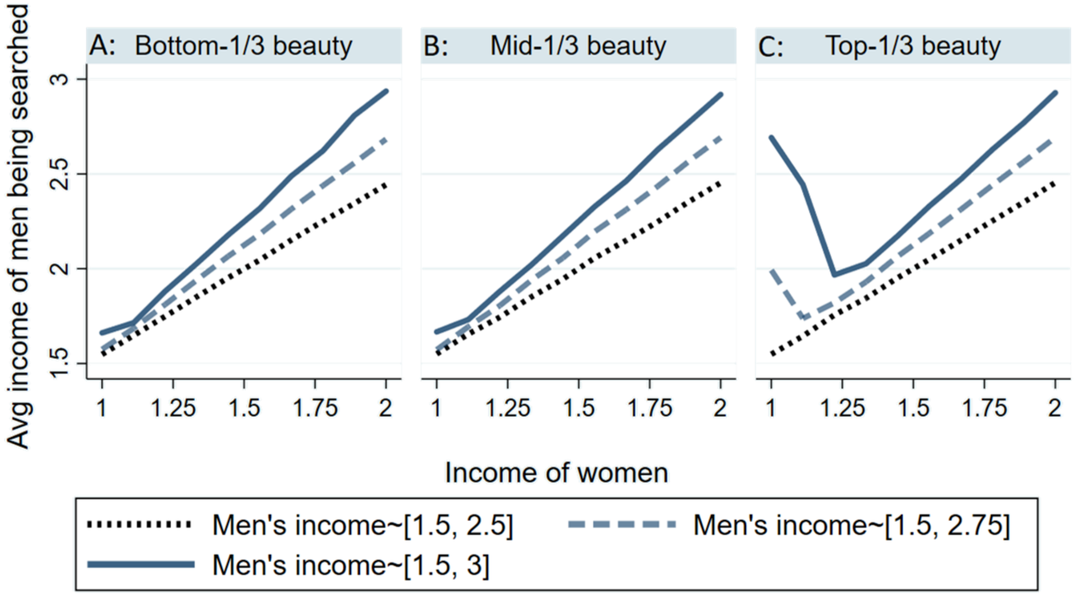
When the sex ratio increases to 1.1, the separation between the dashed lines and the dotted line in panels A and B shows the effect of beautiful low-income women's entry into the high-income men submarket. There is no separation in panel C. The decrease in match probability of high-income women is most 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.

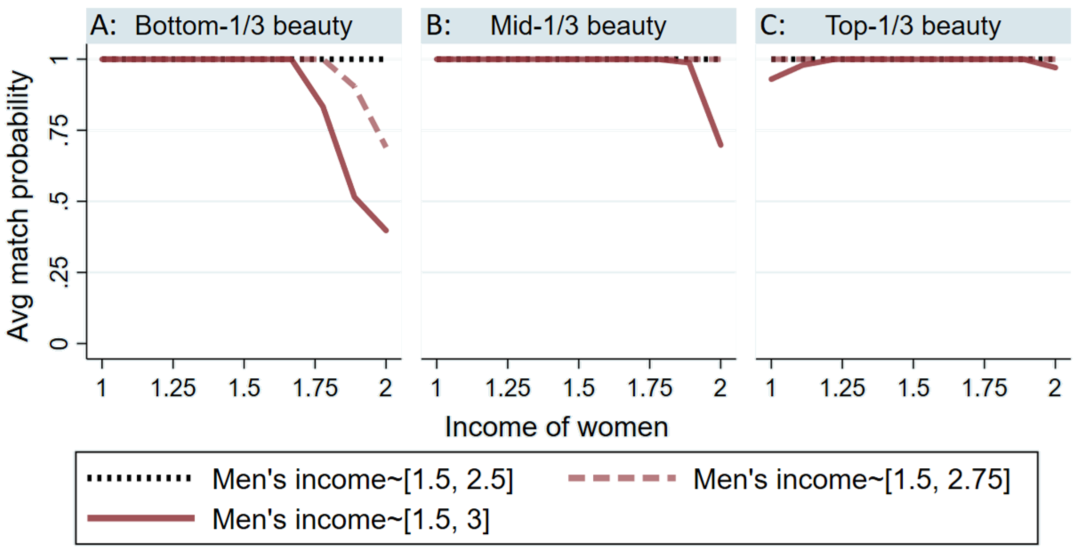
Next, we fix the sex ratio to 1 and conduct two additional simulations in which we increase men's income. Compared to the baseline, the first additional simulation increases the upper bound of men's income distribution from 2.5 to 2.75 and the second one further increases it to 3. We compare them with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

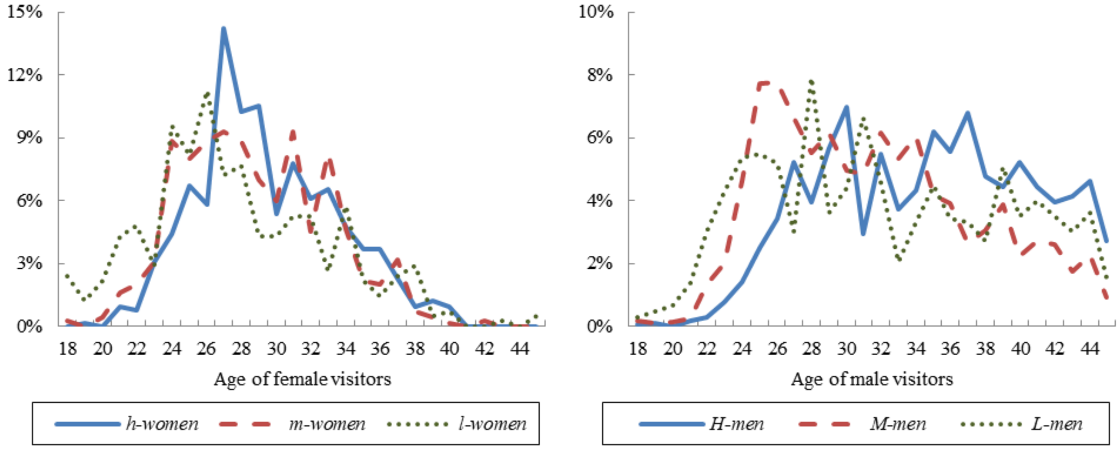
Male	Obs.	Mean	Std. Dev.	Min	Max
Age	5981	33.93	7.580	18	69
Income (1*k* CNY)	5706	10.39	11.03	1	50
Education (years)	5705	15.14	1.689	12	21

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.

References for Appendix

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

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

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.

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

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

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.

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

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

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

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.

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

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*.

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.

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.

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

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. (2019). Directed Search and Competitive Search: A Guided Tour. *Working Paper*.