Terms of Endearment: An Equilibrium Model Of Sex and Matching*

Peter Arcidiacono
Duke University

Andrew Beauchamp
Boston College

Marjorie McElroy
Duke University

May 5, 2011

ABSTRACT: We develop a directed search model of relationship formation which can disentangle male and female preferences for types of partners and for different relationship terms using only a cross-section of observed matches. Individuals direct their search to a particular type of match on the basis of (i) the terms of the relationship, (ii) the type of partner, and (iii) the endogenously determined probability of matching. If men outnumber women, they tend to trade a low probability of a preferred match for a high probability of a less-preferred match; the analogous statement holds for women. Using data from National Longitudinal Study of Adolescent Health we estimate the equilibrium matching model with high school relationships. Variation in gender ratios is used to uncover male and female preferences. Estimates from the structural model match subjective data on whether sex would occur in one's ideal relationship. The equilibrium result shows that some women would ideally not have sex, but do so out of matching concerns; the reverse is true for men.

*We thank Aloysius Siow as well as seminar participants at Boston College, Calgary, Georgetown, Pittsburgh and Princeton. This paper uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu). No direct support was received from grant P01-HD31921 for this analysis.

1 Introduction

With respect to men cheating, "That's a thing that girls let slide, because you have to. ... If you don't let it slide, you don't have a boyfriend (UNC coed)."[1] With respect to the success of his marriage, "... If I had married someone who was more educated or taller than [my wife] Thuy, I don't think she would have been happy here with me (Korean farmer)."[2] These quotes from individuals facing unfavorable gender ratios indicate circumstances in which individuals may sacrifice either their preferred relationship terms or partner type for a higher chance of matching. This paper presents a two-sided model of relationship formation which identifies separate preferences for men and women, enabling the analysis of such trade-offs. Using data on current high school relationships, we present strong evidence that, compared to women, men have a much stronger preference for relationships with sex. Thus, when men are relatively scarce, women agree to sexual relations out of matching concerns.

Disentangling male and female preferences regarding sexual behaviors requires that, ceteris paribus, the extra utility from a given change in the terms of the relationship must differ between men and women. Moreover, as the search behaviors of men and women are rarely observed,[3] we need to be able to identify the separate

[1]Quoted by Williams (2010) in an article discussing social life and relationships at the University of North Carolina, Chapel Hill which has a 40% male and 60% female student body.

[2] Quoted by Onishi (2010) in "Wed to Strangers, Vietnamese Wives Build Korean Lives," the second of two articles describing brokered marriages for rural South Korean men facing a shortage of potential brides.

[3]One exception is data from an on-line matching site collected by Hitsch, Hortacsu and Ariely (2010) and used, in part, to investigate the characteristics that men and

preferences of men and women from data on existing matches.

In order to disentangle male and female preferences from observed matches, we need to specify two features of the market. First is how partners are assigned. In contrast to Choo and Siow (2006), Siow (2009) and Fox (2010), the assignment of partners in our model is not efficient: not everyone who desires to be matched will be. Instead, we allow for search frictions whereby the ex ante yield from searching for a partner can only be known probabilistically. Further, individuals are able to target their search towards individuals with particular characteristics (directed search), potentially trading off the probability of finding a partner against partner quality. Second is how behaviors in relationships are determined. Here we again take a targeted search approach, somewhat analogous to wage posting models such as Burdett and Mortensen (1998). In wage posting models, workers know ahead of time what wage they will receive at a particular employer should they be hired. Our model incorporates directed search on both partner characteristics and relationship terms.

The supply of men and women of different characteristics, coupled with the directed search model allows to uncover differences in male and female preferences. In particular, we rely on the competitive behavior of men and women when searching for a partner. The main idea is that when men outnumber women, we tend to observe relationships characterized by what women want and conversely if women outnumber

women value in a potential partner. They used individual email contact decisions to estimate separate preferences for men and women.

men.[4,5] Men and women target their searches not only based upon the characteristics of the partner but also on the terms of the relationship. For example, a man may choose to search for a woman of a specific race where the relationship would include sex. With the terms of the relationship specified up front, utility is non-transferable. The probability of successfully finding a match then depends upon the number of searchers on each side of the market looking for each combination of race and relationship terms. Searchers face a trade-off between having a low probability of matching under their preferred relationship terms and a higher probability of matching under less-preferred terms. For a large class of constant elasticity of substitution matching functions, we show that, as the gender ratio becomes more unfavorable, the individual becomes more likely to sacrifice relationship terms for a higher match probability.

The advantages of our modeling approach are three-fold. First, by linking choices over partners with outcomes, in equilibrium we are able to weight the different sex ratios appropriately. Standard practice is to use only one sex-ratio when looking at the relationships between sex ratios and outcomes. But the relevant sex ratios are

[4]This fundamental idea has a long pedigree in the literature on intra-household allocations. McElroy and Horney (1981) and McElroy (1990) pointed to the gender ratio in the remarriage market as one member of a class of shifters (EEPs) for the bargaining powers of spouses and thereby intra-household allocations. Chiappori (1992) (and elsewhere) suggested these same shifters (rechristened as "distribution factors") to study intra-household welfare.

[5]Many others have examined the influence of gender ratios on outcomes. See Angrist (2002) for a detailed review of the influence of gender ratios on marriage, labor supply, and child welfare among many others.

determined by equilibrium forces. Second, by allowing preferences over both partner characteristics and what happens in the relationship (in this case age and race of the partner plus whether sex occurs), we are able to capture the tradeoffs made across the two. For example, increasing the number of senior boys favors women. This may, however, still result in relationships that are more sexual because of female preferences for senior boys despite females preferring not to have sex. Finally, by working in a non-transferable framework with partner selection, we are able to identify preferences from data on outcomes. In the transferable case, if a particular outcome is affected by the sex ratio, it is unclear how preferences are affected by the sex ratio because individuals may be making transfers that are not observed by the researcher, particularly in the case where the outcomes of relevance are discrete.

We estimate the model using data from the National Longitudinal Study of Adolescent Health (Add Health). These data contain information on the universe of students at particular U.S. high schools in 1995 as well as answers to detailed questions about relationships for a subset of the students. The model is estimated assuming that individuals are able to target their search towards opposite-sex partners of a particular grade and race as well as to specify whether or not sex will occur in the relationship.

Not surprisingly, estimates of this structural model show that men value sexual relationships relatively more than women. By simulating choices in the absence of matching concerns, we find that 36% of women and 63% of men would prefer to be in a sexual, as opposed to a nonsexual, relationship. These counterfactual choices bear a striking resemblance to subjective reports by students found in Add Health. There, 36% of women and 59% of men responded that sex would be a part of their ideal relationship. Hence, our structural model, while estimated on observed matches, is able to back out preferences for sex that are remarkably close to the self reports,

providing some credence to both the self-reported data and our structural estimates. Taken together, they provide strong evidence that, relative to women, men prefer relationships that include sex.

More importantly, these estimates imply that matching concerns lead some women to have sex, not because they prefer this, but because they were willing to trade off relationship terms for a higher probability of matching. With differing preferences across men and women, observed changes in sexual behavior may then indicate transfers in welfare from one gender to the other.

The rest of this paper proceeds as follows. The next section presents the Add Health data on high school relationships. Section 3 lays out a two-sided model of targeted search and matching, relates the matching function to special cases found in the literature, establishes the existence of equilibrium, and how the gender ratio affects the probability of matching. Section 4 describes the maximum likelihood estimator. Section 5 presents the resulting estimates and shows how the structural model can back out preferences in the absence of competitive effects, demonstrating how the model matches self-reported preferences on a number of dimensions. Section 6 shows our results are robust to different assumptions regarding searching outside the school as well as choosing not to search. Section 7 offers an exploration of what our results imply about female welfare beyond the teen sex setting.

2 Data and Descriptive Characteristics

We use data from Wave I of the National Longitudinal Survey of Adolescent Health.[6] The data include an in-school survey of almost 90,000 seventh to twelfth grade students at a randomly sampled set of 80 communities across the United States.[7] Attempts were made to have as many students as possible from each school fill out the survey during a school day. Questions consist mainly of individual data like age, race, and grade, with limited information on academics, extra-curricular activities and risky behavior. We use this sample to construct school level aggregates by observable characteristics, grade and race, which serve as inputs in calculating gender ratios.

The Add Health data also includes a sample of students who were administered a more detailed survey, the in-home sample. The in-home sample includes a detailed survey about relationship histories and sexual behaviors. The relationship histories include both what happens within the relationships as well as characteristics of the partner such as race and grade. A natural problem in this survey design is the issue of what constitutes a relationship to respondents, particularly when men and women may define relationships differently. Here we follow the Add Health definition that a "relationship" referred to from here on, consists of all the following (i) as holding hands, (ii) kissing, and (iii) saying "I love you." This definition results in the most

[6]The survey of adolescents in the United States was organized through the Carolina Population Center and data were collected in four waves, in 1994-95, 1995-96, 2001-02 and 2008.

[7]A school pair, consisting of a high school and a randomly selected feeder school (middle school or junior high school from the same district) were taken from each community.

symmetric distribution of responses within schools and allows for the most data in the survey to be accessed.[8] The panel-structure of relationships also allows us to determine whether they had sex prior to the current partnership.

We restrict attention to schools which enroll both men and women. A sample of ongoing relationships showed 54.66% of partners met in the same school, the next closest avenue of matching was mutual friends, accounting for 24.08% of matches.[9] Since the focus here will be on a cross section of the matching distribution, we count only current relationships among partners who attend the same school.

The Add Health data is nationally representative at the level of the school and is drawn from all types of schools. We focus on respondents who are in the 9th through 12th grades.[10] Schools for whom we observe fewer than 10 students in the detailed interviews are dropped. We drop one all boys school, one vocational education school for high school dropouts, and we drop six schools without meaningful numbers of 9th-graders.[11] After these adjustments, our sample contains 74 schools, with 11,273

[8]Applying this definition 48.6% of ongoing in-school relationships came from men and 51.4% from women. With perfect reporting and agreement over the definition we would see parity.

[9]The data on where individuals met is only available for a subsample of relationships. The other avenues were 9.38% prior friends, and less than 5% in their neighborhood, place of worship, and casual acquaintances each.

[10]The in-home sample is drawn from schools with different grades: 73% of schools have grades 9-12, 11% have grades 7-12, and 13% had other combinations of grades(e.g. K-12). Finally 1.4% are drawn from a junior and senior high school which are distinct schools.

[11]These schools on average had around 300 students in each of the grades 10-12,

individuals.[12]

Our focus is on matches within a school so those matched with someone outside of the school are dropped. We examine the issue of outside the school matches in more detail in section 6. The sample size removing these individuals falls from 11,273 to 7,915. Since we only observe matches for the in-home sample, we must take into account the fraction of the in-school sample who would also likely be in a relationship outside of the school. We do this by assuming the fraction of the in-home sample in matches outside of the school— conditional on gender, race, grade, and sexual history—matches the fraction of the in-school sample that are in matches outside of the school.

In theory, men and women should report roughly the same number of relationships but in practice this is not the case. Given that we observe double reporting in these data, i.e. men are asked to report their matches within a school and so are women, we can see these differences. Men reported 584 matches where sex occurred and 487 matches where sex did not occur, while women reported 478 matches with sex and 447 matches without sex. To deal with this misreporting we use information about matches reported by women. We drop all men and create matched-male observations from data reported by women,[13] giving us a sample of matched men and women

but on average 9 students in the 9th grade. The Add Health sampling design only probabilistically included the most relevant junior high or middle school for a high school,the relevant 9th grade observations these six schools were not sampled, but rather a small "feeder" school.

[12]From 14840 students between grades 9-12, we drop: individuals missing sample weights(1124), schools discussed above(2337), schools wither fewer than 10 reported students(106).

[13] Partner's past sexual experiences were not reported, so we estimate this proba-

and unmatched women. To get the number of unmatched men we take the original male observations and subtract the number of matched men. All these operations are done at the gender-school-grade-race level. This procedure amounts to treating male reporting of own grade and race as truthful, and female reporting on matches as truthful. The result is 934 current in-school matches, among 7,961 individuals from 74 schools.[14]

2.1 Descriptive Statistics

Table 1 reports descriptive statistics for the sample, prior to the procedure outlined above. Once those matched outside the school are removed, a little under 25% of the sample are in relationship with half of those relationships involving sex. Included in the descriptive statistics is whether the individual has had sex prior to their current relationship. This variable was created from reports of the full relationship history and takes on a value of one if the person has had sex in the past with someone besides their current partner. Men report significantly higher rates of prior sex. There are a disproportionate fraction of the sample in lower grades. This is due to individuals in higher grades being more likely to match outside of the school.

Some direct information on gender differences in preferences for sex can be found from questions that were asked of the in-home sample. Individuals were asked about whether they would want a romantic relationship over the next year and what sorts of things would happen in the relationship. Included in the questions were whether

bility conditional in grade and race from the male reports.

[14]The numbers 7,961 and 7,915 do not match precisely because for some male school-grade-race subgroups the in-home sample interviewed fewer individuals than females matched with, expanding the number of individuals we observe.

the ideal relationship would include having sex.[15] Table 2 shows elicited preferences over sex and relationships overall and by grade. Comparing Table 1 to Table 2, more individuals prefer having relationships than do, suggesting significant search frictions.

While preferences for relationships are the same for both men and women, preferences for sex are not. While 59% of men would prefer to have sex, the fraction of women who prefer to have sex is only 36%. Preferences for sex rise with age. Even with this rise, comparing the sex preferences for women of a particular grade with the sex preferences for men of another grade shows stronger male preferences for sex with one exception: 12th grade women have stronger preference for sex than 9th grade men. Note from Table 1 that half of current relationships entail sex, which is higher than the self-reported preferences for women averaged over any grade, even conditional on wanting a relationship. This suggests the possibility that women may be sacrificing what they want in order to form relationships.

Whether sacrifices over the terms of the relationship are made may in part be dictated by the characteristics of the partner. Individuals may be willing to take more undesirable relationship terms when the partner is more desirable. We now turn to characteristics of the partner, focusing in particular on grade and race. Table 3 shows the share of relationships for each possible male/female grade combination. The most common matches are among individuals in the same grade. Same grade matches make up over 42% of all matches. The six combinations of an older man with a younger woman also make up a large fraction of observations at 40%, leaving only 18% of matches for women with younger men. While matched women are evenly

[15] Add Health responses to questions regarding ever having sex are very similar to the NLSY97 (cf. Arcidiacono, Khwaja and Ouyang (2009)): beginning at a twelfth grade sex participation rate in the low 60% range, and falling roughly 10% per grade.

distributed across grades, older men are substantially more likely to be matched than younger men. Even though 9th grade men outnumber 12th grade men by almost three to two, there are 2.5 times more matched 12th grade men than 9th grade men. These results point towards younger women and older men being more desirable and hence they may have more control over the terms of the relationship.

Table 4 shows the patterns of cross-racial matching. As can be seen from the diagonal elements of the table, the vast majority of matches—over 86%—are same-race matches. Each race is substantially more likely to be in a relationship with someone of their own race than another race. In the set of minorities, Hispanic students date outside their race most often, followed by those in the other category (who are predominantly Asian), and then blacks. Hispanic and black men see much higher probabilities of matching with other races than their female counterparts while the reverse is true for whites and those in the other category.[16] Although not shown here, Hispanic and black men were both more likely to have sex with white female partners conditional on matching, than with partners of their own race. This finding also suggests race-specific gender ratio differences may affect the likelihood of these matches having sex. Black men and, to a lesser extent, black women, make up a larger fraction of matches than they do a percentage of the population.

[16] Other studies have used multiple sources to quantify which races and genders do and do not engage in inter-racial dating: Lee and Edmonston (2005) offer many descriptives using U.S. Census data to track inter-racial marriage over the last 40 years. The census shows a clear pattern with black men and Asian women marrying outside their race far more than black women and Asian men. Qian (1997) reports that white men marry most frequently Asians, Hispanics and lastly blacks.

2.2 Subjective Preferences by Type

How do these variations in partner desirability affect whether sex occurs in the relationship? Table 5 compares elicited preferences for sex versus whether sex occurs in the relationship. The first column shows the percentage of women and men of different grades and races who, conditional on being in a relationship, had sex. The second column then shows the corresponding fraction that wanted to have sex conditional on either wanting a relationship or wanting to have sex without a relationship.[17] While Table 5 shows that across the board women are having sex more than they would like and men are having sex less than they would like, some groups must compromise their preferences more than others. In particular, conditional on being in a relationship, the fraction of 12th grade women who are having sex is much higher than the fraction who would like to be having sex. This is in contrast to 9th grade women whose preferences for sex are similar to what actually occurs. For men, it is the 12th grade men whose preferences align with what actually happens in their relationships while 9th grade men in relationships are having substantially less sex than they would like. Hence, older men and younger women benefit from being more desirable by having relationships which align more with their stated preferences while older women and younger men must sacrifice their preferences over what happens in a relationship in order to successfully match.

[17] Calculated by taking the number of individuals who said that they would ideally have sex and dividing by the number of individuals who said they either wanted a relationship or wanted to have sex without a relationship.

2.3 Gender Ratio Variation

Given evidence that certain characteristics influence whether one's preferences will align with what happens in a relationship, the supply of these characteristics may also have an effect on the terms of the relationship. When men, and in particular older men, are in short supply, women may need to sacrifice their preferences more in order to successfully match. We examine how gender ratios vary across schools in Table 6, paying particular attention to the gender ratios for whites and blacks by grade. Each cell in Table 6 gives the percentage of female students for each grade-race pairing.[18] Table 6 shows that there is a substantial amount of variation in the percentage of female students, particularly among blacks.[19] Breaking out the percentage female along different dimensions (race, and grade-race groupings) spreads the initially condensed distribution.[20]

The bottom panel of Table 6 splits the sample of women by the mean own-grade percent female and calculates the fraction having sex conditional on being in a relationship. The basic split shows a correlation between the percent female and the chances of having sex. Splitting the sample based on whether a woman has had sex in the past shows that higher percent female is more positively correlated with current sex. Arcidiacono, Khwaja and Ouyang (2009) establish that having sex in the

[18]A minimum of 5 observations from the race or grade-race pair is required for a school to enter Table 6

[19]This dispersion is even more pronounced for Hispanic and Asian students.

[20]The populations have been scaled down by one minus the empirical conditional probability of matching outside the school for each age-race-gender-school group. Below we inflate this fraction removed, reflecting that P^{out}_{match} need not equal one (i.e. some non-matched searched outside but failed to match.

past substantially increases the likelihood of having sex in the future. Hence, when women outnumber men, we see what appears to be more substitution toward sex among those women who have had sex in the past.

To formalize this notion more, Table 7 presents probit estimates of the probability of sex conditional on matching for women from the within-school sample of matches. Rather than the own-grade gender ratio, the gender ratio used for each individual is the percent-female in the grade-race pair where the match occurred. For a woman matching with a 12th grade white male, this variable is the fraction of 12th grade whites who are female.[21] Given the high likelihood of individuals matching in their own grade-race pair, this variable serves as a crude measure of the relevant gender ratio a *partner* faces. This is what matters for bargaining: a partners' outside options. The results from the reduced form are clear, with increases in the outside options for male partners increasing the chances of sex occurring. As the second and third column indicate, these results hold regardless of whether observed school characteristics or school effects are included.

Nonetheless, the estimates in 7 do not account for the fact that the grade-race pair individuals matched with is itself a choice. The structural model outlined below specifically accounts for the endogeneity of associated with the choice of partner characteristics. The model will implicitly control for all the factors included in the middle column of 7.[22]

[21] Matched GR $= \sum_m^M 1\{j = m\}\%female_j$ where j is the partners' type and $\%female_j$ is the percent of j-type individuals who are female.

[22] Since controlling for school fixed effects does not significantly alter the reduced form results, we do not estimate the structural model wtih school fixed effects. Doing so would change the number of parameters from 22 to 95.

3 Model

We analyze the tradeoffs amongst three fundamental sources of expected utility from searching for a partner: the type of partner (race, grade), the terms of the relationship (sex/no sex) and the probability of success (matching) in the search for that type and those terms. A searcher knows in advance his utility from type and terms. And, he can target a less-preferred combination of type and terms in order to have a higher probability of matching. It is this fundamental tradeoff that distinguishes our model from others. At its core, our model embeds search and the attendant risk of not matching into a static model. In that sense, it is analogous to the wage posting models in which workers choose between a high wage firm with a low probability of matching, or a low wage firm with a high probability of matching.

In order to disentangle male and female preferences from data on matches, we propose a two-sided search model with non-transferable utility. We consider only opposite-sex one-to-one matching.[23] We categorize each male as a type m where $m \in \{1, 2, \ldots, M\}$. Similarly, we each woman is given a type w where $w \in \{1, 2, \ldots, W\}$ elements. An individual's type denotes some collection of characteristics such as age, grade, race, or attractiveness. For males (females) there are then W (M) types of mates. Let im indicate the i-th member of type m.

We index the possible terms of the relationship by $r \in \{1, ..., R\}$. The possible terms could include not having sex, having sex with protection, etc. We model search as being completely directed: men and women are able to target their search on both

[23] Only 2% of the sample reported concurrent sexual matches and 1% reported concurrent relationships, though clearly some reporting problems exist. We proceed in modeling one-to-one matching given the complexity of modeling multiple matches and the first order importance of the main reported match for preferences.

the characteristics of the partner as well as the terms of the relationship. Each man (woman) then makes a discrete choice to search in one of $M \times R$ ($W \times R$) markets, resulting in $M \times W \times R$ types of matches.

Search is then modeled as a one-shot game: there are no dynamics in the model. Individuals first decide in which market to search. Couples are matched with the probability of matching depending on the number of searchers on both sides of the market.

3.1 Individuals

An individual's expected utility for searching in a particular market depends upon three factors:

1. the probability of matching in the market where the probability of a m-type man matching with a w-type woman in an r-type relationship is P_m^{wr},

2. a deterministic portion of utility conditional on matching given by μ_m^{wr} for a m-type man,

3. and an individual-specific preference term ϵ_{im}^{wr}.[24]

Note that the only individual-specific part of expected utility are the ϵ_{im}^{wr}'s. Further, the ϵ_{im}^{wr}'s are known to the individual before making their decision: there is no match-specific component beyond what occurs through the observed characteristics of the partner and the terms of the relationship. Hence, the only uncertainty from the individual's perspective is their probability of finding a match. Finally, note that the probability of matching is only affected by male and female type and relationship

[24] The corresponding terms for women are P_w^{mr}, μ_w^{mr}, and ϵ_{iw}^{mr}.

type: all males of type m searching in the w, r market have the same probability of matching.

We normalize the utility of not matching to zero. Expected utility from searching in a particular market is then the probability of matching in the market times the utility conditional on matching. We specify the functional form of the utility such that expected utility for a m-type man searching for a w-type woman of relationship type r as:

$$E\left(U_{im}^{wr}\right) = P_m^{wr} \cdot e^{\mu_m^{wr} + \epsilon_{im}^{wr}} \tag{1}$$

We assume that the ϵ_{im}^{wr}'s are i.i.d. Type I extreme value errors and are unknown to the econometrician. Taking logs yields:

$$\ln\left(E\left(U_{im}^{wr}\right)\right) = \mu_m^{wr} + \ln\left(P_m^{wr}\right) + \epsilon_{im}^{wr} \tag{2}$$

Individual i of type m then chooses to search for a woman of type w under relationship terms r, $d_{im} = \{w, r\}$, when:

$$\{w, r\} = \arg\max_{w', r'} \mu_m^{w'r'} + \ln\left(P_m^{w'r'}\right) + \epsilon_{im}^{w'r'}$$

We treat the ϵ_{im}^{wr}'s as observed only to the individual: it is not known to the other participants in the market. The probability of a m-type man searching for a w-type woman in an r-type relationship, ϕ_m^{wr} then follows a multinomial logit form:

$$\Pr(w, r|m) = \phi_m^{wr} = \frac{\exp\left(\mu_m^{wr} + \ln\left(P_m^{wr}\right)\right)}{\sum_{w'} \sum_{r'} \exp\left(\mu_m^{w'r'} + \ln\left(P_m^{w'r'}\right)\right)} \tag{3}$$

where the μ's for one of the markets for both men and women must be normalized to zero.

3.2 Matching

We now specify the matching process. The matching process essentially a production function, taking as inputs the number searching men and the number of searching women in each market and giving the number of matches in each market as output. We parameterize the number of matches, X, in market $\{m, w, r\}$ as depending upon the number of m-type men and w-type women searching in the market. Let N_m and N_w indicate the number of m-type men and number of w-type women. Recall that ϕ_m^{wr} and ϕ_w^{mr} give the probability of m-type men and w-type women who search in market $\{m, w, r\}$ which are also the market shares of searching men and women. The number of matches in market $\{m, w, r\}$ is then given by:

$$X_{mwr} = A\left[(\phi_m^{wr} N_m)^\rho + (\phi_w^{mr} N_w)^\rho\right]^{\frac{1}{\rho}} \tag{4}$$

where ρ determines the elasticity of substitution ($\sigma = \frac{1}{1-\rho}$), and A measures search frictions. When $\rho \to 0$ CES becomes Cobb-Douglas and $\rho \to -\infty$ CES becomes Leontief. Note that $X_{mwr} = X_{wmr}$ for all m, w, and r.

Under the assumption that all m-type men searching in the same market have the same probabilities of matching, P_m^{wr} is given by:

$$\begin{aligned} P_m^{wr} &= \frac{X_{mwr}}{\phi_m^{wr} N_m} \\ &= \frac{A\left[(\phi_m^{wr} N_m)^\rho + (\phi_w^{mr} N_w)^\rho\right]^{\frac{1}{\rho}}}{\phi_m^{wr} N_m} \\ &= A\left[1 + \left(\frac{\phi_w^{mr} N_w}{\phi_m^{wr} N_m}\right)^\rho\right]^{\frac{1}{\rho}} \end{aligned} \tag{5}$$

The log of this term then enters into the multinomial logit probabilities of searching in particular markets and captures the influence of the gender ratio on market search

decisions.

3.3 Equilibrium

The probabilities of searching in a particular market, the ϕ's, give the share of a particular set individuals who will search in a particular market. These ϕ's also affect the probabilities of matching, the P's. We rewrite equation (6) to make the dependence of P_m^{wr} on ϕ_m^{wr} and ϕ_w^{mr} explicit:

$$\phi_m^{wr} = \frac{\exp\left(\mu_m^{wr} + \ln\left[P_m^{wr}\left(\phi_m^{wr}, \phi_w^{mr}\right)\right]\right)}{\sum_{w'}\sum_{r'}\exp\left(\mu_w^{w'r'} + \ln\left[P_m^{w'r'}\left(\phi_m^{w'r'}, \phi_{w'}^{mr'}\right)\right]\right)} \tag{6}$$

Since the market shares must sum to one for both men and women, equilibrium in our model is characterized by stacking the $(M-1)\times(W-1)\times(R-1)$ shares and solving for the fixed point. Since ϕ is a continuous mapping on a compact space, Brouwer's fixed point theorem guarantees that an equilibrium exists.[25] It is trivial to demonstrate the ex ante efficiency of the equilibrium: moving any player from his chosen equilibrium sub-market to another reduces his expected utility and therefore cannot be a Pareto move.

3.4 Implications of Changing the Gender Ratio

Given our utility specification and matching process, we now turn to how changing the gender ratio leads to changes in the probabilities of choosing particular markets. To begin, consider two markets that include w type women and m type men but where the relationship terms in the two markets are given by r and r' respectively. Now, fix the search probabilities, the ϕ's, and increase the number of m-type men. We

[25] As in macro models of the labor market, uniqueness depends on constant returns to scale of the matching function; see Petrongolo and Pissarides (2001).

can then see which of the two relationship markets become relatively more attractive for men and women respectively. We then show how the search probabilities must adjust in equilibrium.

Denoting G_w^m as the ratio N_m/N_w, Proposition 3.1 shows the relationship between the gender ratio and search behavior, with the proof in the appendix.

Proposition 3.1. *If $\rho < 0$ and $\mu_w^{mr'} - \mu_w^{mr} > \mu_m^{wr'} - \mu_m^{wr}$ then the following hold:*

(a) $\frac{\phi_w^{mr}}{\phi_m^{wr}} < \frac{\phi_w^{mr'}}{\phi_m^{wr'}}$

(b) $P_w^{mr} > P_w^{mr'}$ *and* $P_m^{wr} < P_m^{wr'}$

(c) Both $\frac{\partial \phi_w^{mr'}/\phi_w^{mr}}{\partial G_w^m} > 0$ *and* $\frac{\partial \phi_m^{wr'}/\phi_m^{wr}}{\partial G_w^m} > 0$

In Proposition 3.1 the average preference of a $\{m, w\}$ pair is such that the women in the pair have a stronger preference for terms r' over r than their male counterparts. The first two claims are intuitive. Claim (a) states that this relative preference by women for r' over r translates into search behavior: in equilibrium, the ratio of search probabilities for r' versus r must result in women searching more in r' relative to men. These differential search probabilities then translate into match probabilities. Since women are relatively more likely than men to search in r', female match probabilities must be lower in r' than in r, with the reverse holding for men, claim (b).

The key result for the empirical work is claim (c). As the gender ratio moves such that men become relatively more abundant, both men and women increase their relative search probabilities in the market where women have a relative preference, in this case r'. The result falls directly out of the elasticity of substitution. Namely, for elasticities of substitution less than one and conditional on ϕ, the lack of substitutability between men and women in the market implies that, as G_w^m increases, larger changes in match probabilities for men (women) will occur where men are relatively

more (less) abundant. The elasticity of the probability of matching with respect to G_w^m for men and women in the $\{m, w, r\}$ market conditional on ϕ are given by:

$$\begin{aligned}
\left.\frac{\partial \ln\left(P_w^{mr}\right)}{\partial \ln(G_w^m)}\right|_{\phi} &= \left[\left(\frac{\phi_w^{mr}}{\phi_m^{wr} G_w^m}\right)^{\rho} + 1\right]^{-1} \\
\left.\frac{\partial \ln\left(P_m^{wr}\right)}{\partial \ln(G_w^m)}\right|_{\phi} &= -\left[\left(\frac{\phi_m^{wr} G_w^m}{\phi_w^{mr}}\right)^{\rho} + 1\right]^{-1}
\end{aligned}$$

With $\rho < 0$, as the ratio of male to female search probabilities $(\phi_m^{wr}/\phi_w^{mr})$ increases, the magnitude of the elasticity falls for females and rises for males. With fixed search probabilities, increasing the number of men relative to women makes the market where women have a relative preference more attractive for both sexes, resulting in shifts by both men and women to that market in equilibrium.

To illustrate this point, consider the Leontief case where ρ moves toward negative infinity $(\sigma \to 0)$. The matching function is given by:

$$X_{mwr} = A \min\left\{\phi_m^{wr} N_m, \phi_w^{mr} N_w\right\}$$

The number of matches is given by whichever side of the market has fewer searchers. Hence, the gender ratio has an extreme effect on the group that is in the majority in a particular market, with no effect on the group in the minority. On the other extreme is the Cobb Douglas case where $\rho \to 0$, implying the elasticity of substitution, σ, is one. The elasticity given in Equation (4) is 0.5 in this case, regardless of the values either the search probabilities or the gender ratio. Hence, if gender ratios do affect relationship terms then this is evidence that the matching function is not Cobb Douglas.

4 Estimation

Having discussed the trends in the data and the modeling approach, we now turn to integrating the data and the model for estimation. Types of men and women are defined at the grade/race level as suggested by the clear differences in matching patterns across race and grade. Second, we classify relationships as one of two types: those that are having sex and those that are not. An individual is defined as being in a relationship without sex if the person meets the standards described previously (holding hands, etc.). An individual is classified as having a relationship with sex if the individual is currently having sex, regardless of his relationship status. With two types of relationships, four grades, and four races, there are then thirty-two markets. The next two subsections put structure on the utility function for purposes of estimation and then show how to form the likelihood function given the constraints posed by the data.

4.1 Utility

Rather than having separate μ's (utilities) for every type of relationship, we put some structure on the utility function. Denote the grade associated with an m-type man as $G_m \in \{1, 2, 3, 4\}$. When a man searches for an w-type woman, the grade of the partner is PG_w. Similarly, $R_m \in \{1, 2, 3, 4\}$ gives the race of an m-type man with the corresponding race of the potential w-type partner given by PG_w. We specify the utility of a non-sexual relationship as a function of the partner's grade and race as well as whether the partner is in the same grade as the searching individual, $SG_{mw} = I(G_m = PG_w)$ where I is the indicator function, and the same race, $SR_{mw} = I(R_m = PR_m)$. Denoting searching in the no-sex market by $r = 1$, we formulate the deterministic part of utility for men and women matching in the no-sex

market as:

$$\mu_{mw1} = \alpha_1 SG_{mw} + \alpha_2 PG_w + \sum_{j=1}^{4} I(PR_w = j)\left[\alpha_{3j} SR_{mw} + \alpha_{4j}\right] \quad (7)$$

$$\mu_{wm1} = \alpha_1 SG_{mw} + \alpha_5 PG_m + \sum_{j=1}^{4} I(PR_m = j)\left[\alpha_{3j} SR_{mw} + \alpha_{6j}\right] \quad (8)$$

where the intercept of a non-sexual relationship is normalized to zero.

To economize on parameters, this specification sets the extra utility associated with being in the same grade or being of the same race to be the same for men and women. The effect of partner grade and race, however, is allowed to vary by gender. The specification is set such that certain race/gender combinations may be more desirable than other race/gender combinations. Further, the strength of same preferences may differ across races which is captured by allowing α_3 to vary by race.[26]

The utility of sexual relationships takes the utility of non-sexual relationships and adds an intercept as well as effects for partner grade and whether the individual is in 9th grade, $G_m = 1$.[27] We furthermore allow whether the ith man of type m has engaged in sex in the past, $PS_{iw} \in \{0, 1\}$, to affect the current utility of sex. Note that we are not specifying that partners have preferences for individuals who have had sex in the past but rather those who have had sex in the past have preferences for sex now. Hence, the types m and w do not include past sex. Denoting searching in the sex market by $r = 2$, we specify the deterministic part of utility for men and

[26]Due to limited observations, we only estimate two coefficients on same race: one for blacks and one for everyone else.

[27]By tenth grade nearly everyone in the sample has transitioned through puberty.

women matching in the sex market as:

$$\mu_{mw2}(PS_{im}) = \mu_{mw1} + \alpha_7 + \alpha_8 PS_{im} + \alpha_9 PG_w + \alpha_{10} I(G_m = 1) \quad (9)$$

$$\mu_{wm2}(PS_{iw}) = \mu_{wm1} + \alpha_{11} + \alpha_8 PS_{iw} + \alpha_{12} PG_m + \alpha_{10} I(G_m = 1) \quad (10)$$

Although men an women may differ in their preferences for sex and the effect of partner grade on the utility of sex, the effect of past sex is constrained to be the same for men and women.[28]

4.2 Forming the likelihood function

We do not observe all matches but only those in the in-home data set. However, we do observe gender, grade, and race for the population of students at each school. By inferring population moments of past sex from the in-home sample, we can construct the choice probabilities for the entire school from the in-home sample. We take the relationships as defined by the women in the Add Health. However, we still need to incorporate the search decisions of the men. We take these from the women as well: the number of males that do not match are given by the number of males in the in-home sample minus the number of men who were reported to match by the women in the in-home sample. Hence, if we observe 100 white males in the 12th grade in the in-home sample at a particular school and the women in the in-home sample reported 25 matches with 12th grade white males, then the 25 white males would be assigned to the various match categories while the other 75 would be assigned as not matching. This assumes that the ratios of men and women are roughly the same in

[28] Allowing coefficients for past sex to vary by gender and other characteristics did not improve model fit.

the in-home sample as in the in-school sample.

The parameters that need to be estimated include those of the utility function and the parameters of the matching function, ρ and A. Denote θ, as the set $\{\alpha, \rho, A\}$. Denote N as the set of students of each type broken out by the fraction of each that has had prior sex. Hence, N contains 64 elements where each element refers to a gender, grade, race, and past sex combination. Denote $y_{iw} = 1$ if the ith woman of type w was in a current relationship (or having sex) at the time of the survey and is zero otherwise. The woman is then considered matched if $y_{iw} = 1$. Note that d_{iw}, the woman's search decision, is observed only if the woman was matched. Hence, we need to integrate out over the search decision for those who are not matched. The log likelihood for the ith woman of type w is then given by:

$$\begin{aligned} L_{iw}(\theta) &= I(y_{iw} = 1)\left[\sum_m \sum_r I(d_{iw} = \{m, r\})\left(\ln\left[\phi_w^{mr}(\theta, N, PS_{iw})\right] + \ln\left[P_w^{mr}(\theta, N)\right]\right)\right] \\ &\quad + I(y_{iw} = 0)\ln\left[\sum_m \sum_r \phi_w^{mr}(\theta, N, PS_{iw}) \times (1 - P_w^{mr}(\theta, N))\right] \end{aligned} \tag{11}$$

Note that the probability of matching is not affected by past sex except through the search probabilities.

Since we are pulling matched men not from the in-home sample directly but rather from questions asked of the woman about her partner, the likelihood for men is more complicated. In particular, we have to integrate out over whether the men have had prior sex to form the unconditional probability of the outcome. Let π_{mk} indicate the proportion of type m men who were in prior sex state k.[29] Integrating out over the

[29]This proportion is taken as the cell mean of the fraction reporting past sex in the in-home sample within each school and m-subgroup. If data were limited, individuals' probability of past sex is estimated across schools within each m-subgroup.

prior sex state leads to the following log likelihood contribution for ith man of type m:

$$L_{im}(\theta) =$$
$$I(y_{im}=1)\left[\sum_{w}\sum_{r} I(d_{im}=\{w,r\})\left(\ln\left[\sum_{k=0}^{1}\pi_{mk}\phi_m^{wr}(\theta,N,k)\right]+\ln\left[P_m^{wr}(\theta,N)\right]\right)\right]$$
$$+\, I(y_{im}=0)\ln\left[\sum_{k=0}^{1}\sum_{w}\sum_{r}\pi_{mk}\phi_m^{wr}(\theta,N,k)\times(1-P_m^{wr}(\theta,N))\right]$$

where the sum over k is taken over the possible prior sex states.

All of the likelihoods described so far were for a generic school. Denote the schools in the data by $s \in \{1,\ldots,S\}$. Summing the log likelihoods over all the possible m types and w types at each school s implies that the parameters can be estimated using:

$$\hat{\theta} = \arg\max_{\theta}\left(\sum_{s}\sum_{m}\sum_{i=1}^{N_m^s} L_{im}^s(\theta)\right)+\left(\sum_{s}\sum_{w}\sum_{i=1}^{N_w^s} L_{iw}^s(\theta)\right)$$

where a fixed point in the search probabilities is solved at each iteration.

5 Results

The estimates of the structural model are presented in Table 8. Key to disentangling male and female preferences given observed matches is the effect of the different gender ratios on the search decisions. These gender ratios manifest themselves through their effect on the probability of matching. The parameters of the matching function, ρ and A, are identified through variation in matches across schools with different gender ratios and the overall match rate respectively. Estimates of ρ are significant and negative, ruling out the Cobb-Douglas matching model and ensuring that gender ratios do affect the likelihood of observing particular matches. The estimates of ρ indicate the elasticity of substitution in match production of 0.690 with a standard error of 0.086.

The first panel of Table 8 shows how sex is valued above and beyond the relationship itself. Consistent with the elicited preferences in Table 2, males on average have stronger preferences for sex than females. Partner grade is measured here as 0 through 3, with zero corresponding to freshmen and three to seniors. For males, but not for females, grade of the partner influences the utility of sex with men preferring to have sex with older partners. Those who have had sex in the past also have a much stronger preference for sex in the present.

The second panel shows how partner characteristics affect the value of a relationship. Here we see that women prefer to be matched with older men and that this preference is stronger than the preference by men to have sex with older women rather than younger women. In contrast, men only prefer older women when the relationship is accompanied by sex. Individuals also prefer to be matched with those in the same grade and minorities prefer to match with one another. Consistent with Table 4, same race preferences are particularly strong for blacks. The estimates show that cross-race matches are more attractive with whites than with other races. However, the relative preferences for males and females of particular races match those in the prior literature. Namely, women prefer black men more than men prefer black women while men prefer women of other races, a category dominated by Asians, more than women prefer men of other races.

In Table 9 we report the equilibrium match probabilities for whites across grades and depending on whether the individual searchers are male and female.[30] The first column P_m^w captures the probability of a man in grade nine matching with women of various grades. As expected, freshmen males see the lowest match probabilities of any group, and those probabilities decline with female grade. The comparable numbers for women are the intersections of the rows with the P_w^m columns (the second column within each cell). Ninth grade women see significantly higher probabilities of matching relative to ninth grade males. Note that

[30]Focusing on whites removes problems from seeing too few individuals in the school.

in almost every cell in the table, the maximum number for women (under the P_w^m columns) is when a woman looks for a sex partner, and the maximum number for men (under the P_m^w columns) is when a man looks for a relationship without sex. This pattern of matching is due to the preference for sex among men, and captures the trade-off between preferred relationship terms and higher match probabilities.

Between the general equilibrium effects and the non-linear nature of the specification, the magnitudes of the utility coefficients are difficult to interpret. However, we can use the coefficient estimates to back out the fraction of men and women who prefer sex to no sex absent concerns about matching. Namely, we can turn off the effects of the probability of matching and see what choices would have been made in the absence of having to compete for partners. Backing out male and female preferences for sex in this way yields estimates of the fraction of each group who prefer to have sex. We then compare the model estimated preferences to the stated preferences discussed in Table 5.

Table 10 shows that the model does a remarkably good job of matching the elicited preferences for sex given that the elicited measures were not used in estimation. The fit is particularly good for women where the elicited preferences show 36.3% of women prefer sex to 35.9% of women predicted by the model, while for men the model predicts 62.5% compared to the elicited 59.1%. The model-predicted preferences for men and women are very close across races and for individuals with no prior sex. The model over-predicts interest in sex for 10th grade men, and both men and women who have had sex in the past.

We next show how male and female probabilities differ due to matching concerns. Table 11 shows differences between male and female search probabilities both with and without allowing the probabilities of matching to affect the search decision. Taking into account the probability of matching, male probabilities of searching in the set of markets that include sex are thirteen percentage points higher than females. Without matching concerns, there is a twenty-six percentage point difference. This means both sides of the market are sacrificing their preferred relationship terms to increase their chances of matching.

6 Robustness Checks

In this section we relax assumptions on who searches outside the school and whether all individuals search, showing that the general pattern of our results regarding differences in preferences between males and females are robust to alternative assumptions. Previously we removed individuals matched with someone outside of the school. This assumption is consistent with either a) individuals searched and matched in the outside market in a first stage or b) the probability of matching being one in the outside market. To see b), note that only those who matched in the outside market will have searched in the outside market. Denote 0 as the outside market. The probability of an m-type man searching in the $\{m, w, r\}$ market is:

$$\Pr(w, r|m) = \frac{\exp\left(\mu_m^{wr} + \ln\left(P_m^{wr}\right)\right)}{\sum_{w'}\sum_{r'}\exp\left(\mu_m^{w'r'} + \ln\left(P_m^{w'r'}\right)\right) + \exp(\mu_m^0)} \tag{12}$$

Using the independence of irrelevant alternatives (IIA) property, note this probability conditional on searching in the school is:

$$\Pr(w, r|m, d \neq 0) = \phi_m^{wr} = \frac{\exp\left(\mu_m^{wr} + \ln\left(P_m^{wr}\right)\right)}{\sum_{w'}\sum_{r'}\exp\left(\mu_m^{w'r'} + \ln\left(P_m^{w'r'}\right)\right)}, \tag{13}$$

which is the probability in (6) that we used to form our likelihoods, assuming everyone who did not match in the outside market searched within the school.

When the probability of matching in the outside market is not one, the IIA property still holds, but we no longer have a good measure of the number of searchers in the outside market. We allow for the probability of matching in the outside market to be less than one by assuming that match rates are the same for all those who search outside the school. Namely, we know the characteristics of those who matched in the outside market. If 10% of freshmen girls and 16% of senior girls match in the outside market, for a candidate match rate of 80%, the implied freshmen girl search rate would 12.5% in the outside market with the corresponding search rates for senior girls being 20%. We then scale down the number

of non-matched individuals with particular characteristics given the implied outside-market search rates to determine the number of women searching within the school. For unmatched individual's in the sample we randomly assign status as outside-searchers, based on their type-school-specific probability of searching outside, which combines data on matching outside with an assumption on the match rate. We then drop these outside-searchers.

More formally, denote N_m^0 as the number of m-type men who matched in the outside market. N_m was defined as the number of searching men of type m in the school and was formed as the number of m-type men minus N_m^0. With the probability of matching in the outside market given by P_m^0, instead of forming N_m by subtracting off N_m^0 for the population, we subtract off N_m^0/P_m^0. Note that the utilities of searching in the outside market are allowed to vary conditional on school and characteristics. The only restriction is that, conditional on searching in the outside market, the match rate is the same. We need this in order to characterize the number of those searching inside the school.[31]

Results for the estimated model under different match rate assumptions are presented in Table 12, with the first column repeating the results of the original model. The top rows of Table 12 show the mean choice probability for choosing sex without equilibrium influences. For reasonable match rates, we continue to see the same patterns of women preferring sex relative to men. However, as the match rate falls, ρ moves closer to zero, which is the Cobb-Douglas case.[32]

[31]The details of how we scale down the number of individuals searching within the school are as follows. Given an estimate of the fraction of individuals of a given type at each school F_m^0, F_w^0, we inflate this fraction by one over match rate outside th school. The number of male searchers of a given type is $N_m = N_m(1 - F_m^0(1/P_m^0))$. Throughout we do not observe N_m^0, but estimate it as $N_m^0 = \hat{F}_m^0 N_m$.

[32]We estimated models with lower out-matching fractions, but ρ becomes very small and insignificant erasing the identifying power of the sex ratios for uncovering gender differences in utility.

As we move across the table, the composition of the sample changes. This is because we drop individual's who we randomly assign as outside-searchers, in proportion individuals in their school-type who matched outside. Most importantly this process yields a sample of in-school searchers who are younger. This is reflected in the declining subjective probability of wanting sex for both men and women across the bottom rows of Table 12. Nonetheless, for low probabilities of matching outside the subjective means are still in a tight range. We use these subjectives as a check on the appropriate outside-market match rate. Here we see that the model with outside the school match rate set to one generates model predictions which match the subjective means the closest.

Another assumption made throughout is that all individuals engaged in search. Using data on the history of matching, we relax this assumption by assuming a given fraction of those individuals who never matched (never had sex, or a relationship in the past or present), were uninterested in searching for an opposite sex partner. Assuming some individuals decided in a first stage not search, we again drop individuals randomly from the group who have never matched and re-estimate the model. We do this both with and without the assumption that all unmatched individuals who searched did so in the school (the assumption relaxed in Table 12). The results are presented in Table 13, where we compare fitness using the subjective probabilities. Results from the baseline model are given in the "None"-"$P_m^o = 1$" column.

The subjective means in the lower panel show as we saw in Table 12, as we decrease P_m^o the fraction preferring sex falls. Now as we also remove individuals who never match the fraction preferring sex rises for both men and women, counteracting the fall because the sample is younger under $P_m^o = .66$ than under $P_m^o = 1$. Note however that the net changes are small as we drop more of the never-matched moving across Table 13.

As can be seen in moving across the upper panel of the table, adjusting the sample along these dimensions worsens the fit of the model relative to the subjective probabilities. In the extreme going from $P_{Match}^{out} = 1$ and everyone searches, and $P_{Match}^{out} = 1$ and only

50% of the never matched searched produce a very similar fit for men, but over-predict the fraction of women interested in sex. Moving from $P^{out}_{Match} = .66$ and everyone searches to $P^{out}_{Match} = .66$ and only half the never-matched population searches improves the fit for women but worsens it for men. There is likely some optimal combination of adjustment along these two dimensions which yields model fitness similar to the baseline, but overall the specifications the qualitative pattern of results (e.g. sex preferences, age profiles and cross-race matching) are very similar.

7 Conclusion

The contribution of this paper is two-fold. First, we show how a directed search model can disentangle male and female preferences for different types of relationships using variation in the gender ratio. When the researcher's goal is to understand differences in male and female preferences, directed search provides an alternative to transferable utility models: transferable utility models are difficult to use here since we rarely observe transfers.

Second, we have applied the directed search model to the teen sex and dating market and uncovered male and female differences in preferences for sex. The preferences from the structural model match the self-reported preferences, providing an out-of-sample test for the validity for the approach. That men and women value sex differently suggests that changes in sexual behavior may have different welfare effects for men than for women. Further, when gender ratios tilt such that men become a minority—as is the case on many college campuses—women are more likely to engage in sex conditional on forming a relationship, sacrificing their preferred relationship terms for a higher probability of matching.

Beyond Teen Sex

The result that individuals trade off preferred characteristics of a match for an increased chance of matching touches a number of areas of concern beyond high school matching. The influential theory put forth by Wilson (1987) holds that the rise in out-of-wedlock childbearing among blacks is due to a shortage of marriageable black men. Taking this

idea farther,Willis (1999) modeled an equilibrium in which, with a shortage of men, the marriage market bifurcates into a richer segment in which children are born within marriage and a poorer segment in which men father children by multiple women who bear primary responsibility raising them. That women faced with a shortage of partners respond by only demanding partial responsibility for child-rearing from men is consistent with our findings

Goldin, Katz and Kuziemko (2006) documented that for US cohorts born after 1959, women's college attendance rates exceeded that of men. Our results suggest these changes in the gender ratio should translate into a higher fraction of relationships involving sex on campus and, later in life, women marrying "down" in order to marry. Indeed, Rose (2003) found that for recent cohorts of highly educated women a decline in hypergamy [marrying up] allowed the marriage market to absorb the increased number of educated women.

In France, Abramitzky, Delavande and Vasconcelos (2010) presented difference-in-difference evidence that France's shortage of men following the carnage of World War I resulted in surviving French men marrying into a higher social classes after the war than similar men before the war, and especially so in regions hardest hit by the death toll.

Highly skewed gender ratios are also a common characteristic of developing nations in Africa and Asia and they have become the object of research and policy by several international agencies.[33] Typically western countries have 1050 women per 1000 men. But in India, for example, the provisional 2005 Census shows that, of all states in India, Punjab (not a poor state) had the lowest ratio, 935 women per 1000 men. Within Punjab the district of Ludhaina had only 824 women per thousand men.[34] Such endogenously determined severe imbalances will no doubt greatly impact marriage patterns, dowries, intra-household

[33]For example, following on Sen (1990) famous calculation, the United Nations Population Fund (2007) put the number of "missing" women and girls in Asia alone at almost 163 million and has made understanding the causes, manifestations and consequences an official goal.

[34]Indian Census for 2005, http://www.censusindiamaps.net/IndiaCensus/Map.htm.

distributions, fertility and the well-being of men and women for generations to come.

Stevenson and Wolfers (2009) found that despite the enormous objective gains of women over the last 35 years (education, wages, income, etc.), subjective self-reported measures of well-being indicate that women's wellbeing has declined both relative to men's as well as absolutely. They found this across various data sets, subjective measures of well-being, demographic groups and industrialized countries. Studies such as ours may begin to chip away at such paradoxes.

References

Abramitzky, Ran, Adeline Delavande, and Luis I Vasconcelos, "Marrying Up: The Role of Sex Ratio in Assortative Matching," Stanford University, Working Paper 2010.

Angrist, Josh, "How Do Sex Ratios Affect Marriage And Labor Markets? Evidence From America'S Second Generation," *The Quarterly Journal of Economics*, August 2002, *117* (3), 997–1038.

Arcidiacono, Peter, Ahmed Khwaja, and Lijing Ouyang, "Habit Persistence and Teen Sex: Could Increased Access to Contraception have Unintended Consequences for Teen Pregnancies?," Duke University, Working Paper 2009.

Burdett, Kenneth and Dale T Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 1998, *39* (2), 257–273.

Chiappori, Pierre-Andre, "Collective Labor Supply and Welfare," *Journal of Political Economy*, June 1992, *100* (3), 437–67.

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

Fox, Jeremy, "Identification in Matching Games," *Quantitative Economics*, 2010, *1* (2), 203–254.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko, "The Homecoming of American College Women: The Reversal of the College Gender Gap," *Journal of Economic Perspectives*, Fall 2006, *20* (4), 133–156.

Hitsch, Guenter J., A. Hortacsu, and Dan Ariely, "Matching and Sorting in Online Dating," *American Economic Review*, March 2010, *100* (1), 130–163.

Lee, S. and B. Edmonston, "New Marriages, New Families: U.S. Racial and Hispanic Intermarriage," *Population Bulletin*, 2005, *60* (2), 3–36.

McElroy, Marjorie B, "The Empirical Content of Nash-Bargained Household Behavior," *Journal of Human Resources*, 1990, *25* (4), 559–583.

_ **and Mary Jean Horney**, "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Consumer Demand," *International Economic Review*, 1981, *22* (2), 333–349.

Onishi, Norimitsu, "Wed to Strangers, Vietnamese Wives Build Korean Lives," *New York Times*, March 2010.

Petrongolo, Barbara and Christopher A. Pissarides, "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, June 2001, *39* (2), 390–431.

Qian, Zhenchao, "Breaking the racial barriers: Variations in interracial marriage between 1980-1990," *Demography*, 1997, *34* (2), 263.

Sen, Amartya, "More Than 100 Million Women Are Missing," *New York Review of Books*, December 1990, pp. 61–66.

Siow, Aloysius, "Testing Becker.s Theory of Positive Assortative Matching," University of Toronto, Working Paper 2009.

Stevenson, Betsey and Justin Wolfers, "The Paradox of Declining Female Happiness," *American Economic Journal: Economic Policy*, August 2009, *1* (2), 190–225.

Williams, Alex, "The new math on campus," *New York Times*, February 2010.

Willis, Robert J., "A Theory of Out-of-Wedlock Childbearing," *Journal of Political Economy*, December 1999, *107* (S6), S33–29.

Wilson, William J., *The Truly Disadvantaged*, University of Chicago Press, 1987.

8 Appendix

Proof of Proposition 3.1 The proof of claim (a) follows from manipulating the definition of the search probability. Assuming $\mu_w^{mr'} - \mu_w^{mr} > \mu_m^{wr'} - \mu_m^{wr}$, we can add the log-match probability for each combination to both sides in the following way:

$$\begin{aligned} &\mu_w^{mr'} + \log(P_w^{mr'}) - \mu_w^{mr} - \log(P_w^{mr}) + \log(P_m^{wr'}) - \log(P_m^{wr}) > \\ &\quad \mu_m^{wr'} + \log(P_m^{wr'}) - \mu_m^{wr} - \log(P_m^{wr}) + \log(P_w^{mr'}) - \log(P_w^{mr}). \end{aligned} \tag{14}$$

Exponentiating both sides gives us a ratio of choice probabilities and match probabilities because the choice probabilities share the same denominator:

$$\frac{e^{\mu_w^{mr'} + \log(P_w^{mr'})}}{e^{\mu_w^{mr} + \log(P_w^{mr})}} e^{\log(P_m^{wr'}) - \log(P_m^{wr})} > \frac{e^{\mu_m^{wr'} + \log(P_m^{wr'})}}{e^{\mu_m^{wr} + \log(P_m^{wr})}} e^{\log(P_w^{mr'}) - \log(P_w^{mr})} \tag{15}$$

or:

$$\frac{\phi_w^{mr'}}{\phi_w^{mr}} \frac{P_m^{wr'}}{P_w^{mr'}} > \frac{\phi_m^{wr'}}{\phi_m^{wr}} \frac{P_m^{wr}}{P_w^{mr}}. \tag{16}$$

Now note the ratio of match probabilities can be expressed as:

$$\frac{P_m^{wr'}}{P_w^{mr'}} = \frac{\left[1 + \left(\frac{\phi_w^{mr'} N_w}{\phi_m^{wr'} N_m}\right)^{\rho}\right]^{1/\rho}}{\left[1 + \left(\frac{\phi_m^{wr'} N_m}{\phi_w^{mr'} N_w}\right)^{\rho}\right]^{1/\rho}} = \frac{\left[\left(\frac{\phi_w^{mr'} N_w}{\phi_m^{wr'} N_m}\right)^{\rho} + \left(\frac{\phi_m^{mr'} N_w}{\phi_m^{wr'} N_m}\right)^{\rho}\right]^{1/\rho}}{\left[\left(\frac{\phi_w^{mr'} N_w}{\phi_w^{mr'} N_w}\right)^{\rho} + \left(\frac{\phi_m^{mr'} N_m}{\phi_w^{mr'} N_w}\right)^{\rho}\right]^{1/\rho}}$$

which by canceling the numerators inside both matching functions simplifies to:

$$\frac{P_m^{wr'}}{P_w^{mr'}} = \left[\left(\frac{\phi_w^{mr'} N_w}{\phi_m^{wr'} N_m}\right)^{\rho}\right]^{1/\rho}.$$

Imposing the same equality in the r-market, and substituting into the inequality we have the following:

$$\frac{\phi_w^{mr'}}{\phi_w^{mr}}\left[\left(\frac{\phi_w^{mr'} N_w}{\phi_m^{wr'} N_m}\right)^{\rho}\right]^{1/\rho} > \frac{\phi_m^{wr'}}{\phi_m^{wr}}\left[\left(\frac{\phi_w^{mr} N_w}{\phi_m^{wr} N_m}\right)^{\rho}\right]^{1/\rho}, \tag{18}$$

which further simplifies to:

$$\left(\frac{\phi_w^{mr'}}{\phi_w^{mr}}\right)^2 > \left(\frac{\phi_m^{wr'}}{\phi_m^{wr}}\right)^2, \tag{19}$$

and claim (a) follows since the choice probabilities are always positive.

Claim (b) follows from claim (a) and $\rho < 0$. Given claim (a) we have:

$$\frac{\phi_w^{mr}}{\phi_m^{wr}} < \frac{\phi_w^{mr'}}{\phi_m^{wr'}} \tag{20}$$

multiplying both sides by N_w/N_m and raising both sides to the $1/\rho$ power flips the inequality:

$$\left(\frac{\phi_w^{mr} N_w}{\phi_m^{wr} N_m}\right)^{1/\rho} > \left(\frac{\phi_w^{mr'} N_w}{\phi_m^{wr'} N_m}\right)^{1/\rho} \tag{21}$$

adding one to both sides and raising both to the power ρ switches the inequality once more and we have:

$$\left[1 + \left(\frac{\phi_w^{mr} N_w}{\phi_m^{wr} N_m}\right)^{1/\rho}\right]^{\rho} < \left[1 + \left(\frac{\phi_w^{mr'} N_w}{\phi_m^{wr'} N_m}\right)^{1/\rho}\right]^{\rho} \tag{22}$$

which is the definition of $P_m^{wr} < P_m^{wr'}$. Beginning with the inequality between the ratio of choice probabilities with female choice probabilities in the denominator delivers the result for female match probabilities.

To evaluate claim (c), we use the implicit function theorem coupled with Cramer's rule. We show the case when there is only one type of man and one type of woman with two relationship types, r and r'. For ease of notation, we then denote $G = G_r^m$. Our proof,

however, holds in the general case due to the independence of irrelevant alternatives associated with the Type I extreme value errors. Note that the definitions of search probabilities imply that, in equilibrium the log odds ratios for women satisfy:

$$\ln(\phi_w^{mr'})-\ln(1-\phi_w^{mr'}) \equiv \ln(\mu_w^{mr'})-\ln(\mu_w^{mr})+\ln\left[1+\left(\frac{\phi_m^{wr'}G}{\phi_w^{mr'}}\right)^{\rho}\right]-\ln\left[1+\left(\frac{(1-\phi_m^{wr'})G}{1-\phi_w^{mr'}}\right)^{\rho}\right] \tag{23}$$

Now, define $F_1(\phi_w^{mr'},\phi_m^{wr'},G)$ and $F_2(\phi_w^{mr'},\phi_m^{wr'},G)$ based on the identity in 23 and the corresponding expression for men, respectively:

$$\begin{aligned} F_1(\phi_w^{mr'},\phi_m^{wr'},G) &\equiv \ln(\phi_w^{mr'})-\ln(1-\phi_w^{mr'})-\ln(\mu_w^{mr'})+\ln(\mu_w^{mr}) \\ &\quad -\ln\left[1+\left(\frac{\phi_m^{wr'}G}{\phi_w^{mr'}}\right)^{\rho}\right]+\ln\left[1+\left(\frac{(1-\phi_m^{wr'})G}{1-\phi_w^{mr'}}\right)^{\rho}\right] \end{aligned} \tag{24}$$

$$\begin{aligned} F_2(\phi_w^{mr'},\phi_m^{wr'},G) &\equiv \ln(\phi_m^{wr'})-\ln(1-\phi_m^{wr'})-\ln(\mu_m^{wr'})+\ln(\mu_m^{wr}) \\ &\quad -\ln\left[1+\left(\frac{\phi_w^{mr'}}{\phi_m^{wr'}G}\right)^{\rho}\right]+\ln\left[1+\left(\frac{1-\phi_w^{mr'}}{(1-\phi_m^{wr'})G}\right)^{\rho}\right] \end{aligned} \tag{25}$$

which can equivalently be expressed as:

$$\begin{aligned} F_1(\phi_w^{mr'},\phi_m^{wr'},G) &\equiv 2\ln(\phi_w^{mr'})-2\ln(1-\phi_w^{mr'})-\ln(\mu_w^{mr'})+\ln(\mu_w^{mr}) \\ &\quad -\ln\left[\left(\phi_w^{mr'}\right)^{\rho}+\left(\phi_m^{wr'}G\right)^{\rho}\right]+\ln\left[(1-\phi_w^{mr})^{\rho}+\left((1-\phi_m^{wr'})G\right)^{\rho}\right] \end{aligned} \tag{26}$$

$$\begin{aligned} F_2(\phi_w^{mr'},\phi_m^{wr'},G) &\equiv 2\ln(\phi_m^{wr'})-2\ln(1-\phi_m^{wr'})-\ln(\mu_m^{wr'})+\ln(\mu_m^{wr}) \\ &\quad -\ln\left[\left(\phi_m^{wr'}\right)^{\rho}+\left(\frac{\phi_w^{mr'}}{G}\right)^{\rho}\right]+\ln\left[(1-\phi_m^{wr'})^{\rho}+\left(\frac{1-\phi_w^{mr'}}{G}\right)^{\rho}\right] \end{aligned} \tag{27}$$

Taking the total derivative of the identities imply from the implicit function theorem that the following holds:

$$\begin{bmatrix} \frac{\partial F_1}{\partial \phi_w^{mr'}} & \frac{\partial F_1}{\partial \phi_m^{wr'}} \\ \frac{\partial F_2}{\partial \phi_w^{mr'}} & \frac{\partial F_2}{\partial \phi_m^{wr'}} \end{bmatrix} \begin{bmatrix} \frac{\partial \phi_w^{mr'}}{\partial G} \\ \frac{\partial \phi_m^{wr'}}{\partial G} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial G} \\ \frac{\partial F_2}{\partial G} \end{bmatrix}$$

which can be expressed as:

$$\begin{bmatrix} \frac{\partial \phi_w^{mr'}}{\partial G} \\ \frac{\partial \phi_m^{wr'}}{\partial G} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial \phi_w^{mr'}} & \frac{\partial F_1}{\partial \phi_m^{wr'}} \\ \frac{\partial F_2}{\partial \phi_w^{mr'}} & \frac{\partial F_2}{\partial \phi_m^{wr'}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial G} \\ \frac{\partial F_2}{\partial G} \end{bmatrix}$$

The partial derivatives of F_1 and F_2 with respect to G can be expressed as:

$$\frac{\partial F_1}{\partial G} = G^{-1} \left(\left[\left(\frac{1-\phi_w^{mr'}}{(1-\phi_m^{wr'})G} \right)^{\rho} + 1 \right]^{-1} - \left[\left(\frac{\phi_w^{mr'}}{\phi_m^{wr'}G} \right)^{\rho} + 1 \right]^{-1} \right) \tag{28}$$

$$\frac{\partial F_2}{\partial G} = G^{-1} \left(\left[\left(\frac{\phi_m^{wr'}G}{\phi_w^{mr'}} \right)^{\rho} + 1 \right]^{-1} - \left[\left(\frac{(1-\phi_m^{wr'})G}{1-\phi_w^{mr'}} \right)^{\rho} + 1 \right]^{-1} \right) \tag{29}$$

$$\tag{30}$$

Since $\rho < 0$ and (by claim (a)) $\frac{1-\phi_w^{mr'}}{1-\phi_m^{wr'}} = \frac{\phi_w^{mr}}{\phi_m^{wr}} < \frac{\phi_w^{mr'}}{\phi_m^{wr'}}$, both $\partial F_1/\partial G$ and $\partial F_2/\partial G$ are both greater than zero.

The partial derivatives of F_1 and F_2 with respect to $\phi_w^{mr'}$ and $\phi_m^{wr'}$ can be expressed as:

$$\frac{\partial F_1}{\partial \phi_w^{mr'}} = \frac{(\phi_w^{mr'})^{\rho} + 2(\phi_m^{wr'}G)^{\rho}}{\phi_w^{mr'}\left((\phi_w^{mr'})^{\rho} + (\phi_m^{wr'}G)^{\rho}\right)} + \frac{(1-\phi_w^{mr'})^{\rho} + 2((1-\phi_m^{wr'})G)^{\rho}}{(1-\phi_w^{mr'})\left((1-\phi_w^{mr'})^{\rho} + ((1-\phi_w^{mr'})G)^{\rho}\right)} > 0 \tag{31}$$

$$\frac{\partial F_1}{\partial \phi_m^{wr'}} = -\frac{(\phi_m^{wr'}G)^{\rho}}{\phi_m^{wr'}\left((\phi_w^{mr'})^{\rho} + (\phi_m^{wr'}G)^{\rho}\right)} - \frac{((1-\phi_m^{wr'})G)^{\rho}}{(1-\phi_m^{wr'})\left((1-\phi_w^{mr'})^{\rho} + ((1-\phi_w^{mr'})G)^{\rho}\right)} < 0 \tag{32}$$

$$\frac{\partial F_2}{\partial \phi_w^{mr'}} = -\frac{(\phi_w^{mr'})^{\rho})}{\phi_w^{mr'}\left((\phi_w^{mr'})^{\rho} - (\phi_m^{wr'}G)^{\rho}\right)} - \frac{(1-\phi_w^{mr'})^{\rho}}{(1-\phi_w^{mr'})\left((1-\phi_w^{mr'})^{\rho} + ((1-\phi_w^{mr'})G)^{\rho}\right)} < 0 \tag{33}$$

$$\frac{\partial F_2}{\partial \phi_m^{wr'}} = \frac{2(\phi_w^{mr'})^{\rho} + (\phi_m^{wr'}G)^{\rho}}{\phi_m^{wr'}\left((\phi_w^{mr'})^{\rho} + (\phi_m^{wr'}G)^{\rho}\right)} + \frac{2(1-\phi_w^{mr'})^{\rho} + ((1-\phi_m^{wr'})G)^{\rho}}{(1-\phi_m^{wr'})\left((1-\phi_w^{mr'})^{\rho} + ((1-\phi_w^{mr'})G)^{\rho}\right)} > 0 \tag{34}$$

Appealing to Cramer's rule,

$$\frac{\partial \phi_w^{mr'}}{\partial G} = \frac{\frac{\partial F_1}{\partial G}\frac{\partial F_2}{\partial \phi_m^{wr'}} - \frac{\partial F_1}{\partial \phi_m^{wr'}}\frac{\partial F_2}{\partial G}}{\frac{\partial F_1}{\partial \phi_w^{mr'}}\frac{\partial F_2}{\partial \phi_m^{wr'}} - \frac{\partial F_1}{\partial \phi_m^{wr'}}\frac{\partial F_2}{\partial \phi_w^{mr'}}} \tag{35}$$

$$\frac{\partial \phi_w^{mr'}}{\partial G} = \frac{\frac{\partial F_2}{\partial G}\frac{\partial F_1}{\partial \phi_w^{mr'}} - \frac{\partial F_2}{\partial \phi_w^{mr'}}\frac{\partial F_1}{\partial G}}{\frac{\partial F_1}{\partial \phi_w^{mr'}}\frac{\partial F_2}{\partial \phi_m^{wr'}} - \frac{\partial F_1}{\partial \phi_m^{wr'}}\frac{\partial F_2}{\partial \phi_w^{mr'}}} \tag{36}$$

In both cases, the numerators are positive. Both have one negative term, $\partial F_1/\partial \phi_m^{wr'}$ and $\partial F_2/\partial \phi_w^{mr'}$ respectively but this term is multiplied by negative 1.

The denominators are the same across the two expressions. The first term is positive but the second term is negative. However, the first term can be written as the negative of the second term plus additional terms. The sign of the denominator is then positive, implying that both expressions are positive as well. QED.

Tables

Table 1: Descriptive Statistics[a]

	Women	Men
Currently		
Matched(sex or relationship)	.243	.261
In a Relationship	.230	.249
Having Sex	.127	.144
Prior sex	.247	.334
Race		
Black	.199	.161
Hispanic	.153	.156
Other	.077	.081
Fraction of Sample		
9th Grade	.294	.284
10th Grade	.273	.273
11th Grade	.235	.245
12th Grade	.199	.198
Mean % Female	.508	.506
N	3847	4068

[a]Based on reporting from both men and women, removing individuals matched outside the school.

Table 2: Stated Preferences[a]

Prefer:	Women	Men
Relationship	.844	.852
Sex	.363	.591
Sex, Own Grade :		
9th	.247	.476
10th	.356	.566
11th	.399	.665
12th	.501	.698
N	3785	4005

[a]Fraction of sample responding "yes," includes current within-school matches and currently unmatched. Number of respondents varies by question, N is the number reporting on whether they wanted to have sex.

Table 3: Cross-Grade Matching Distribution[a]

	Male Grade					
Female Grade	9th	10th	11th	12th	Total	% Sample
9th	**8.6**	6.2	5.5	3.7	24.0	29.4
10th	2.9	**9.1**	8.1	8.2	28.1	27.3
11th	2.1	4.0	**9.4**	8.8	24.3	23.5
12th	0.5	1.5	7.0	**14.6**	23.6	19.9
Total	14.1	20.8	30.0	35.2	100.0	100.0
% Sample	28.3	27.2	24.5	20.1	100.0	

[a]Numbers are the percent of 934 total matches observed in the within school sample.(e.g. 8.6% of the 934 matches were between ninth graders). % Sample refers to percent of the total sample (matched and unmatched) in each group.

Table 4: Cross-Race Matching Distribution[a]

	Male Race					
Female Race:	White	Black	Hispanic	Other	Total	% Sample
White	**56.3**	1.3	3.9	0.8	62.3	57.2
Black	0.6	**19.7**	1.0	0.0	21.3	19.9
Hispanic	2.9	0.5	**8.4**	0.1	11.8	15.3
Other	1.1	0.7	0.8	**1.9**	4.6	7.7
Total	60.9	22.2	14.1	2.9	100.0	100.0
% Sample	59.7	16.3	15.8	8.1	100.0	

[a]Numbers are the percent of 934 total matches observed in the within school sample (e.g. 56.3% of the 934 matches were white-white matches). % Sample refers to percent of the total sample(matched and unmatched) in each race-group.

Table 5: Stated Preferences and Sex[a]

	Female		Male	
	Sex Occurred	Preferred Sex	Sex Occurred	Preferred Sex
9th Grade	.326	.247	.383	.478
10th Grade	.496	.358	.434	.566
11th Grade	.557	.398	.513	.664
12th Grade	.717	.501	.638	.701
White	.522	.357	.510	.571
Black	.598	.398	.607	.711
Hispanic	.421	.375	.448	.627
Other	.432	.298	.500	.445
N	934	3324	934	3723

[a]Sample sizes reflect match distribution and responses to subjective preference questions. Number of observations varies by group, N given is the unconditional set of respondents.

Table 6: Variation in Fraction Female[a]

	Percentile	
% Female by Race-Grade:	.25	.75
Total	.479	.523
White	.479	.536
9th	.464	.545
10th	.464	.554
11th	.434	.538
12th	.473	.570
Black	.464	.546
9th	.424	.591
10th	.462	.621
11th	.444	.586
12th	.444	.636

	Fraction Female Own Grade	
Observed:	< Median	> Above Median
P(Sex|Match)	.490	.551
N	457	477
P(Sex|Match,Female No Past Sex)	.356	.392
N	320	324
P(Sex|Match,Female Past Sex)	.803	.889
N	137	153

[a]Based on a sample of 74 schools. Gender ratios calculated using only those searching within the school. Probability of sex conditional on matching is calculated from only observed in-school matches. Aggregate gender ratio refers to the fraction of the population that is female.

Table 7: Probit of Sex Conditional on Matching[a]

	P(Sex|Match)	P(Sex|Match)	P(Sex|Match)
Matched Gender Ratio	.738*	.946**	1.064**
	(.392)	(.426)	(.538)
Prior Sex	1.599*	1.559**	1.634**
	(.340)	(.338)	(.392)
School Characteristics	No	Yes	No
School Fixed Effects	No	No	Yes
N	933	933	929

[a]Observations in regression are females. Matched Gender Ratio takes the percent female within the partners' grade and race pair. Gender ratios use estimates of the fraction of each observed type searching within the school. Regressions control for race by partner race fixed effects, linear terms for grade and partner grade, prior sex and prior sex interacted with grade and partner grade. School characteristics are percent nonwhite, school level number non-virgin males and females and school size. Standard errors are clustered at the school level. * and ** denote significance at the 90% and 95% level respectively.

Table 8: Model Estimates, In School Matches[a]

	Utility		
Utility from Sex	Male	Female	Common
Intercept (α_7, α_{11})	-.770*	-.9491**	-
	(.4289)	(.414)	
Partner Grade (α_9, α_{12})	.570**	.015	-
	(.155)	(.124)	
Past Sex (α_8)	-	-	2.062**
			(.113)
Own Grade = 9 (α_{10})	-	-	-.843**
			(.206)
Utility from Matching			
Partner Grade (α_2, α_5)	-.026	1.150**	-
	(.087)	(.139)	
Partner Black (α_{4b}, α_{6b})	1.621**	2.142**	-
	(.173)	(.115)	
Partner Hisp. (α_{4h}, α_{6h})	-.745**	-.420**	-
	(.302)	(.105)	
Partner Other (α_{4o}, α_{6o})	-1.709**	-2.989**	-
	(.482)	(.321)	
Same Grade (α_1)	-	-	.685**
			(.081)
Same Race (α_3)	-	-	-.170
			(.830)
Same Race × Black (α_{3b})	-	-	1.875**
			(.848)

Matching Parameters	Coefficient	s.e.
ρ	-.514**	.047
A	.313**	.015
-log(L)	8460.6	
N	7961	

[a]* and ** denote significance at the 10% and 5% levels respectively. Standard errors are in parentheses. Sample includes only non-matched and those reporting current in-school matches. Market populations excludes those currently matched outside the school.

Table 9: Equilibrium Probabilities of Matching: Whites[a]

Female Grade and Relationship:	Male Grade 9		10		11		12	
	P_m^w	P_w^m	P_m^w	P_w^m	P_m^w	P_w^m	P_m^w	P_w^m
9								
sex	.135	.556	.211	.425	.348	.272	.503	.158
no sex	.148	.530	.294	.321	.464	.184	.635	.097
10								
sex	.127	.570	.199	.441	.333	.286	.485	.167
no sex	.120	.586	.249	.373	.409	.222	.575	.121
11								
sex	.088	.658	.141	.531	.252	.367	.394	.226
no sex	.097	.631	.208	.419	.359	.258	.525	.144
12								
sex	.061	.742	.103	.619	.189	.455	.309	.296
no sex	.084	.676	.189	.460	.326	.294	.483	.168

[a] Each cell gives the probability of matching in sex or no sex markets based on an individuals' grade and possible partner grade. P_m^w is the probability of matching for a man looking for a woman, P_w^m is the probability of matching for a woman looking for a man.

Table 10: Stated vs. Predicted Preferences for Sex[a]

Subgroup:	Female Stated	Female Model	Male Stated	Male Model
All	.363	.359	.591	.625
9th Grade	.247	.210	.476	.439
10th Grade	.356	.398	.566	.653
11th Grade	.399	.426	.665	.707
12th Grade	.501	.449	.698	.747
Black(mean across grades)	.398	.399	.711	.699
Hispanic(mean across grades)	.346	.336	.627	.649
Prior Sex	.622	.705	.838	.879
No Prior Sex	.277	.246	.506	.502

[a]Predicted means set the probability of matching to one, giving the average choice probability across schools based only on preferences absent matching concerns.

Table 11: Effect of Matching on Choice of Sex[a]

	No Equilibrium	With Equilibrium
In-School Matches	$P^{sex}_{male} - P^{sex}_{female}$	$P^{sex}_{male} - P^{sex}_{female}$
All	.265	.126
9th Grade	.229	.118
10th Grade	.256	.116
11th Grade	.281	.133
12th Grade	.298	.130

[a]Gives average choice probability differences between men and women based on preference parameter estimates for the No Equilibrium column, and based on preference and matching function estimates in the With Equilibrium column.

Table 12: Robustness Check: P^{out}_{match} [a]

	P^{out}_{match}			
Mean ϕ_{sex}	1.0	.90	.66	.50
Male, No Equilibrium	.625	.807	.809	.806
Female, No Equilibrium	.359	.237	.215	.253
Mean Stated Preference: Sex				
Male	.591	.585	.574	.562
Female	.363	.356	.346	.346

[a]Columns are separate estimations. Columns shrink the number of individuals searching inside the school by the relevant school-grade-race-gender-prior sex specific fraction matching outside the school.

Table 13: Robustness Check: Varying P^{out}_{match} and Non-searchers[a]

	% Never Matched Removed					
	None		25%		50%	
Mean ϕ_{sex}	$P^o_m = 1$	$P^o_m = .66$	$P^o_m = 1$	$P^o_m = .66$	$P^o_m = 1$	$P^o_m = .66$
Male	.625	.809	.619	.791	.581	.729
Female	.359	.215	.372	.254	.469	.341
Stated Preferences						
Male	.591	.574	.603	.595	.629	.616
Female	.363	.346	.372	.369	.396	.388

[a]Each cell re-estimates the model. Some individuals with no history of matching are assumed to not search and are randomly removed. In panel two 25% are removed, in panel three 50%. The number of individuals searching in the school is also reduced in both cases, using the conditional probability of never matching in the past.