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

Peter Arcidiacono  Andrew Beauchamp  Marjorie McElroy
Duke University & NBER  Boston College  Duke University & NBER

May 2013

Abstract: We develop a two-sided directed search model of relationship formation which disentangles male and female preferences over partner characteristics and different relationship terms, using only a cross-section of observed matches. Individuals gain utility from relationships and direct their search for a partner on the basis of (i) the terms of the relationship, (ii) the partners’ characteristics, and (iii) the endogenously determined probability of matching. If men outnumber women, in equilibrium 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 this equilibrium matching model with high school relationships. Variation in gender ratios is used to uncover male and female preferences. Estimates from this 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. Counterfactual simulations show the matching environment black women face is the primary driver of the large differences in sexual activity among white and black women.

*We thank Aloysius Siow as well as seminar participants at Boston College, Boston University, Calgary, Georgetown, NYU, Pittsburgh, Princeton and participants at the Incentives and Matching in Marriage and Dating Markets session at the 2012 AEA Winter Meetings and the 2011 Cowles Conference on Structural Microeconomics. This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). 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 suggest individuals may sacrifice either their preferred relationship terms or partner characteristics 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 for relationship terms 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, we need to be able to identify the separate 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. 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. Our approach is then a middle ground between models where the assignment to partners is efficient\(^3\) and search models where the probability of matching with a partner of a particular type is exogenous, though the choice of whether to accept a match is endogenous.\(^4\)

\(^1\)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.


\(^3\)See Choo and Siow (2006), Siow (2009) and Fox (2010).

Second, we have to specify 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.\(^5\)

The supply of men and women of different characteristics, coupled with the directed search model allows us to uncover differences in male and female preferences. In particular, we rely on the competitive behavior of men and women when searching for a match. The main idea is that when men outnumber women, we tend to observe relationships characterized by what women want and conversely if women outnumber men.\(^6\) Men and women target their searches not only based upon the multidimensional characteristics of the partner but also on the terms of the relationship.\(^8\) 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.\(^9\) 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

---

\(^5\)Dagsvik (2000) develops a theoretical matching model with terms, where agents have preferences over terms and observed match characteristics, and the preference distribution can in principle be backed out from observed aggregate matching patterns. Recent work by Salani and Galichon (2012) addresses matching based on unobserved characteristics.

\(^6\)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 using these same shifters to study intra-household welfare.

\(^7\)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.

\(^8\)Recent empirical work by Dupuy and Galichon (2012) shows sorting in the marriage market is not unidimensional and in individuals trade-off heterogeneous characteristics differently.

\(^9\)While we assume non-cooperative behavior in the teen matching market, recent work by Del Boca and Flinn (2012a) shows non-cooperative behavior (such as the “separate spheres” behavior of Lundberg and Pollak (1993)) can appear within marriage as well. Del Boca and Flinn (2012b) derive an estimator based on matching patterns to distinguish models of non-cooperative household interaction from those with some form of cooperation.
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 gender ratios appropriately. Standard practice is to use only one sex-ratio when looking at the relationships between gender ratios and outcomes. But the relevant gender ratios are 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 when the matching environment changes. For example, increasing the number of senior boys favors women. This may, however, still result in more sexual relationships because the female preference for senior boys overcomes females’ preferring not to have sex. Finally, by working in a non-transferable framework with partner selection, we are able to identify preferences for outcomes. In the transferable case, if the occurrence of a particular outcome is affected by the gender ratio, it is unclear how utility are affected by the gender ratio because individuals may be making transfers 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 in our sample. By simulating choices in the absence of matching concerns, we find that 34.1% of women and 57.4% of men would prefer to be in a sexual, as opposed to a nonsexual, relationship. These counterfactual

---


4
choices bear a striking resemblance to subjective reports by students found in Add Health. There, 33.2% of women and 58.1% 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. 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.

Because we estimate the full structural model, we can use the parameter estimates to perform counterfactual experiments. The data reveal that black females (and males) are substantially more likely to have sex than their white counterparts, conditional on matching. We simulate changing the market black females face in order to understand the sources of this racial gap. We do so in two steps, first examining the impact of blacks facing the same grade-specific gender ratios as whites, and secondly the impact of facing the same distribution of sexually-experienced teens in the school. While changing the gender ratios has a substantial impact on match probabilities, their effect on the probability of sex conditional on matching is smaller. The primary driver behind differences in sex behavior between white women and black women is the strong preferences for own-race matches coupled with black males being substantially more sexually-experienced than white males.

The rest of this paper proceeds as follows. The next section 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 3 presents the Add Health data on high school relationships. 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. We also show our results are robust to different assumptions regarding searching outside the school as well as choosing not to search. Section 6 offers an exploration of what our results imply about female welfare beyond the teen sex setting.
2 Model

We analyze the tradeoffs among 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). 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, we propose a two-sided search model with non-transferable utility and consider only opposite-sex, one-to-one matching.\(^\text{11}\) We categorize each male as a type \(m\) where \(m \in \{1, 2, \ldots, M\}\). Similarly, each woman is given a type \(w\) where \(w \in \{1, 2, \ldots, W\}\) elements. An individual’s type can denote some collection of observed characteristics such as age, grade, race, or attractiveness. For males (females) there are then \(W(M)\) types of mates. Let \(i_m\) indicate the \(i\)-th member of type \(m\).

We index the possible terms of the relationship by \(r \in \{1, \ldots, R\}\). 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 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.

\[^{11}\text{Only 2\% of the sample reported concurrent sexual matches and 1\% reported concurrent relationships, though clearly some reporting problems exist. We proceed in modeling one-to-one matching given the complexity of modeling multiple matches and the first order importance of the main reported match for preferences.}\]
2.1 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 \( P_{wrm} \),

2. a deterministic portion of utility conditional on matching given by \( \mu_{wr} \) for a \( m \)-type man,

3. and an individual-specific preference term \( \epsilon_{wr}^{im} \).\(^{12}\)

Note that the only individual-specific part of expected utility are the \( \epsilon_{im}^{wr} \)'s. Further, the \( \epsilon_{im}^{wr} \)'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 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:

\[
E(U_{im}^{wr}) = P_{wrm} \cdot e^{\mu_{wr}^{im} + \epsilon_{im}^{wr}}
\]  

(1)

Taking logs yields:

\[
\ln(E(U_{im}^{wr})) = \mu_{wr}^{im} + \ln(P_{wrm}) + \epsilon_{im}^{wr},
\]  

(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:

\[
\{w, r\} = \arg \max_{w', r'} \mu_{m}^{w'r'} + \ln(P_{m}^{w'r'}) + \epsilon_{im}^{w'r'}
\]

\(^{12}\)The corresponding terms for women are \( P_{mr}^{wr} \), \( \mu_{mr}^{wr} \), and \( \epsilon_{im}^{mr} \).
We treat the $\epsilon_{imr}$’s as observed only to the individual: only the distribution is known to the other participants in the market. We assume that the $\epsilon_{imr}$’s are i.i.d. Type I extreme value errors and are unknown to the econometrician. In this case, we can estimate the error variance, $\sigma$, as a coefficient on the log probability term, capturing how the probability of matching influences utility. The probability of a $m$-type man searching for a $w$-type woman in an $r$-type relationship, $\phi_{mr}^{w}$, then follows a multinomial logit form:

$$
\Pr(w, r|m) = \phi_{m}^{wr} = \frac{\exp \left( \frac{\mu_{mr}^{w} + \ln(P_{mr}^{w})}{\sigma} \right)}{\sum_{w'} \sum_{r'} \exp \left( \frac{\mu_{m'}^{r'} + \ln(P_{m'}^{r'})}{\sigma} \right)}
$$

(3)

2.2 Matching

We now specify the matching process. The matching process is 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 an 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 overall. Recall that $\phi_{m}^{wr}$ and $\phi_{w}^{mr}$ 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. Thus $\phi_{m}^{wr} N_{m}$ is the number of men of type $m$ searching women of type $w$ on relationship terms $r$. The number of matches in market $\{m, w, r\}$ is then given by:

$$
X_{mwr} = A \left[ \frac{(\phi_{m}^{wr} N_{m})^{\rho}}{2} + \frac{(\phi_{w}^{mr} N_{w})^{\rho}}{2} \right]^{\frac{1}{\rho}}
$$

$$
= A \left[ (\phi_{m}^{wr} N_{m})^{\rho} + (\phi_{w}^{mr} N_{w})^{\rho} \right]^{\frac{1}{\rho}}
$$

(4)

where $\rho$ determines the elasticity of substitution ($\frac{1}{1-\rho}$), and $A$ measures search frictions. When $\rho \to 0$ the CES function becomes Cobb-Douglas, and as $\rho \to -\infty$ the CES function 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, $P_{wr}^m$ is given by:

$$P_{wr}^m = \frac{X_{mwr}}{\phi_{m}^{wr} N_m}$$

$$= A \left[ (\phi_{m}^{wr} N_m)^{\rho} + (\phi_{w}^{mr} N_w)^{\rho} \right]^{\frac{1}{\rho}}$$

$$= A \left[ 1 + \left( \frac{\phi_{w}^{mr} N_w}{\phi_{m}^{wr} N_m} \right)^{\frac{1}{\rho}} \right].$$

(5)

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

### 2.3 Equilibrium

The probabilities of searching in a particular market, the $\phi$’s, give the share of a particular set individuals who will search in a particular market. These $\phi$’s also affect the probabilities of matching, the $P$’s. We rewrite equation (6) to make the dependence of $P_{wr}^m$ on $\phi_{m}^{wr}$ and $\phi_{w}^{mr}$ explicit:

$$\phi_{m}^{wr} = \frac{\exp \left( \frac{\mu_{m}^{wr} + \ln[P_{m}^{wr}(\phi_{m}^{wr}, \phi_{w}^{mr})]}{\sigma} \right)}{\sum_{w'} \sum_{r'} \exp \left( \frac{\mu_{w'}^{r'} + \ln[P_{m}^{wr'}(\phi_{m}^{wr'}, \phi_{w'}^{mr'})]}{\sigma} \right)}$$

(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 $\phi$ is a continuous mapping on a compact space, Brouwer’s fixed point theorem guarantees that an equilibrium exists.\(^{13}\) 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.

\(^{13}\)As in macro models of the labor market, uniqueness depends on constant returns to scale of the matching function; see Petrongolo and Pissarides (2001).
2.4 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 \( \phi \)'s, and increase the number of \( m \)-type men. We 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 \( G_{mw} \) as the ratio \( N_m/N_w \), Proposition 1 shows the relationship between the gender ratio and search behavior, with the proof in Appendix A.

**Proposition 1.** If \( \rho < 0 \) and \( \mu_{wr}^{mr} - \mu_{wr}^{mr'} > \mu_{mr}^{wr} - \mu_{mr}^{wr'} \) then the following hold:

\[
\begin{align*}
(a) & \quad \phi_{mr}^{wr}/\phi_{mr}^{wr'} < \phi_{mr'}^{wr}/\phi_{mr}^{wr'} \\
(b) & \quad P_{mr}^{wr'} > P_{mr}^{wr} \text{ and } P_{wr}^{mr} < P_{wr}^{mr'} \\
(c) & \quad \text{Both } \frac{\partial \phi_{mr}^{wr'/}\phi_{mr}^{wr}}{\partial G_{mw}} > 0 \text{ and } \frac{\partial \phi_{mr}^{wr'/\phi_{mr}^{wr}}}{\partial G_{mw}} > 0
\end{align*}
\]

In Proposition 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 our 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 (\( \rho < 0 \)) and conditional on \( \phi \), the lack of substitutability between men and women in the market implies that, as \( G_{mw} \) 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 $G_m^w$ for men and women in the \{$m, w, r$\} market conditional on $\phi$ are given by:

$$\frac{\partial \ln (P_{mr}^w)}{\partial \ln (G_m^w)} \mid \phi = \left[\left(\frac{\phi_{mr}^w G_m^w}{\phi_{mr}^w G_m^w} + 1\right)^\rho + 1\right]^{-1}$$

$$\frac{\partial \ln (P_{mr}^w)}{\partial \ln (G_m^w)} \mid \phi = -\left[\left(\frac{\phi_{mr}^w G_m^w}{\phi_{mr}^w G_m^w} + 1\right)^\rho + 1\right]^{-1}$$

With $\rho < 0$, as the ratio of male to female search probabilities ($\phi_{mr}^w / \phi_{mr}^w$) 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 $\rho$ moves toward negative infinity ($\sigma \to 0$). The matching function is given by:

$$X_{mwr} = A \min \{\phi_{mr}^w N_m, \phi_{mr}^w N_w\}$$

The number of matches is determined 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, $\sigma$, is one. The implied 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.

3 Data and Descriptive Characteristics

We use data from Wave I of the National Longitudinal Survey of Adolescent Health.\textsuperscript{14} 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.\textsuperscript{15} At-

\textsuperscript{14} 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.

\textsuperscript{15} 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.
tempts 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, allowing us to calculate gender ratios.

The Add Health data also includes a random sample of students who were administered a more detailed in-home survey. The in-home sample includes detailed histories and sexual behaviors. The relationship histories include both what happens within the relationships as well as characteristics of the partner. 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 use the definition that a “relationship” referred to from here on, consists of all the following (i) as holding hands and (ii) kissing. This definition results in the most symmetric distribution of responses within schools and allows for the most data in the survey to be accessed.\textsuperscript{16} The history also allows us to determine whether they had sex with a different partner prior to the current partnership.

We restrict attention to schools which enroll both men and women. A sample of ongoing relationships showed 55\% of partners met in the same school, the next closest avenue of matching was mutual friends, accounting for 24\% of matches.\textsuperscript{17} 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 school-level and is drawn from all types of schools. We focus on respondents who are in the 9th through 12th grades.\textsuperscript{18} 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.\textsuperscript{19} After these adjustments, our sample contains 74 schools, with 11,273

\textsuperscript{16} Applying this definition 48.8\% of ongoing in-school relationships came from women and 51.2\% from men. With perfect reporting and agreement over the definition we would see parity.

\textsuperscript{17} The data on where individuals met is only available for a subsample of relationships. The other avenues were 9\% prior friends, and less than 5\% in their neighborhood, place of worship, and casual acquaintances each.

\textsuperscript{18} 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.

\textsuperscript{19} These schools on average had around 300 students in each of the grades 10-12, but on average 9 students in the 9th grade. The Add Health sampling design only probabilistically included the
Our focus is on matches within a school. Those matched with someone outside of the school are initially dropped, consistent with assuming the outside matching market is frictionless, which we relax in section 6. After removing these individuals, the sample size falls to 7,280. 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 601 matches where sex occurred and 566 matches where sex did not occur, while women reported 502 matches with sex and 545 matches without sex. Unfortunately we cannot link individuals to their partners within the sample, meaning incorporating both sets of reports would double count a subset of matches. To deal with both double counting and misreporting we use information about matches reported by women. This also ensures that our estimated results on differences in male and female preferences for sex are driven not by differences in self-reporting but from differences in the gender ratio.

We drop all men and create matched-male observations from data reported by women, giving us a sample of matched men and women and unmatched women. To get the number of unmatched men we take the original male observations and subtract the number of matched men. This procedure amounts to treating male reporting of own grade and race as truthful, and female reporting on matches as truthful. The result is 1047 current in-school matches, among 7,355 individuals from 74 schools. The numbers 7,280 and 7,355 do not match precisely because for some male school-grade-race
Table 1 reports descriptive statistics for the sample. Roughly 30% of the sample are in current relationship and roughly half of these involve sex. Included in the descriptive statistics is whether the individual has had sex prior to their current relationship, which varies both by gender and race.\textsuperscript{24} Men report significantly higher rates of prior sex than women. Black men and women have had sex more than whites, but the gender gap in past sex is larger for blacks and Hispanics. Current sexual participation is also higher for black males and females than the other race-groups.

\subsection*{3.1 Direct Measures of Preferences}

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 physical events would occur between the partners. Included in the questions were whether the ideal relationship would include having sex.\textsuperscript{25} Table 2 shows elicited preferences over sex and relationships overall, by grade and by race. 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 (roughly 90\% want a relationship as defined), preferences for sex are not. While 57\% of men would prefer to have sex, the fraction of women who prefer to have sex is only 34\%. 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.

\textsuperscript{24}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.

\textsuperscript{25}Add Health responses to questions regarding ever having sex are very similar to the NLSY97 (cf. Arcidiacono, Khwaja and Ouyang (2012)): beginning at a twelfth grade sex participation rate in the low 60\% range, and falling roughly 10\% per grade.
To investigate this further, Table 5 shows the probability of having sex conditional on whether the respondent’s ideal relationship includes sex. The means are presented separately for men and women, and show that women who want to have sex are significantly more likely to have sex than men who want to have sex. Further, women who don’t want to have sex are also significantly more likely to have sex than men who don’t want to have sex. The second row show that these male/female differences hold conditional on being matched: it is not just that women who want sex sort into relationships at a higher rate, they also see their preferences implemented within matches more frequently than similar men. In contrast, women who do not want to have sex see their preferences implemented with matches less frequently than similar men. Finally we also condition on having had sex in past, where we see the difference is largely driven by differences among the sexually inexperienced: in this group women are 14 points more likely to have sex conditional on wanting to have sex.

3.2 The Distribution of Matches

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 41% of all matches. The six combinations of an older man with a younger woman also make up a large fraction of observations at over 40%, leaving less than 20% of matches for women with younger men. While matched women are evenly 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. 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.\footnote{27} 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. Roughly 5\% of black male matches are with white females, similar to the fraction addressed in Wong (2003b), who argues a marriage taboo dramatically influences the frequency of cross-race matching among black males. Black men and, to a lesser extent, black women, also make up a larger fraction of matches than they do a percentage of the population.

3.3 The Gender Ratio and Its Implications for Relationship Terms

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.\footnote{28} Table 6 shows that there is a substantial amount of variation in the percentage of female students, particularly among blacks.\footnote{29} Breaking out the percentage female along different dimensions (race, and grade-race groupings) spreads the initially condensed distribution.\footnote{30}

\footnote{27}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.\footnote{28}A minimum of 5 observations from the race or grade-race pair is required for a school to enter Table 6\footnote{29}This dispersion is even more pronounced for Hispanic and Asian students.\footnote{30}The populations have been scaled down by one minus the estimated conditional probability of matching outside the school for each age-race-gender-school group. Below we inflate the fraction removed, reflecting that \( P_{\text{out}} \) need not equal one (i.e. some non-matched searched outside but
The bottom panel of Table 6 shows the probability of having sex conditional matching conditional on the fraction female being above or below the median. The first two rows show the cases when fraction female is measured using the whole school and then using only those of the same race. In both cases, a higher fraction female is associated with more sex, though the differences are not large. The evidence in section 3.2 showed that the most common matches are between those of the same race and grade so we next consider the percent-female of the same grade-race of the partner. For a woman matching with a 12th grade white male, this variable is the fraction of 12th grade whites who are female. Given the high likelihood of individuals matching in their own grade-race pair, this variable serves as a crude measure of the outside options the partner faces. The final row of Table 6 shows that if the fraction female in the partner’s race-grade combination is above the median the probability of sex in the relationship is more likely: more competition for women implies more sex.

To further investigate the role of competition in determining relationship terms, Table 7 presents marginal effects from a probit estimated on the probability of sex conditional on matching for women. We again use the the fraction female in the partner’s grade-race combination as this gives us our best reduced form measure of the partner’s outside options. The results from the reduced form are clear: increases in the outside options for male partners is associated with a higher probability of having sex. The second column indicates these results are strengthened when we control for some basic characteristics of the school population, with school fixed effects (column 3) further strengthening the results. Increasing partner grade also affects the probability of having sex, even conditional on own grade and prior sex. Since older men appear to be more desirable for women, this suggests women are willing to give on their preferred relationship terms in order to match with a more-preferred partner.

Note that the estimates in Table 7 do not account for the fact that the grade-race pair individuals matched with is itself a choice. The structural model outlined above specifically accounts for the endogeneity of associated with the choice of partner characteristics.

failed to match).
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. 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 and shows how to form the likelihood function given the constraints posed by the data. However, there are three additional issues which arise from our data: (1) the unobserved fraction of each individual type within each school who match outside the school, (2) the unobserved distribution of past sex among men, and (3) unreported partner characteristics. Further, with one independent error draw for each type of partner and relationship term, groups of smaller sizes will attract more attention due to the error draws above and beyond the effect of gender ratios. We discuss how we deal with each of these issues in Appendix B.

4.1 Utility

Rather than having separate $\mu$’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:

\[
\mu_{mw1} = \alpha_1 SG_{mw} + \alpha_2 PG_w + \alpha_3 SR_{mw} + \sum_{j=1}^{4} I(PR_w = j)\alpha_{4j}
\]  
(7)

\[
\mu_{wm1} = \alpha_1 SG_{mw} + \alpha_5 PG_m + \alpha_3 SR_{mw} + \sum_{j=1}^{4} I(PR_m = j)\alpha_{6j}
\]  
(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.

The utility of sexual relationships takes the utility of non-sexual relationships and adds an intercept as well as allowing whether the individual has had sex in the past, \(PS_{iw}\), 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, and it is therefore not targeted. Denoting searching in the sex market by \(r = 2\), we specify the deterministic part of utility for men and women matching in the sex market as:

\[
\mu_{mw2}(PS_{im}) = \mu_{mw1} + \alpha_7 + \alpha_8 PS_{im}
\]  
(9)

\[
\mu_{wm2}(PS_{iw}) = \mu_{wm1} + \alpha_{11} + \alpha_8 PS_{iw}
\]  
(10)

Although men and women may differ in their preferences for sex, the effect of past sex is constrained to be the same for men and women.\(^{31}\)

4.2 Forming the likelihood function

We do not observe all matches but only those in the in-home sample. 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

\(^{31}\)Allowing coefficients for past sex to vary by gender generates a problem with identification since we must integrate out over the probability that each male has had sex in the past because it is unobserved from female reporting.
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 matches while the other 75 would be designated as not matching.

The parameters that need to be estimated include those of the utility function and the parameters of the matching function, \( \rho \) and \( A \). Denote \( \theta \), as the set \( \{ \alpha, \rho, A, \sigma \} \). 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 \( i \)th 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 \( i \)th woman of type \( w \) is then given by:

\[
L_{iw}(\theta) = \begin{cases} 
I(y_{iw} = 1) \left[ \sum_m \sum_r I(d_{iw} = \{m, r\}) \left( \ln [\phi_{mrw}(\theta, N, PS_{iw})] + \ln [P_{mrw}(\theta, N)] \right) \right] \\
+ I(y_{iw} = 0) \ln \left[ \sum_m \sum_r \phi_{mrw}(\theta, N, PS_{iw}) \times (1 - P_{mrw}(\theta, N)) \right] 
\end{cases}
\]

(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 \( \pi_{mk} \) indicate the proportion of type \( m \) men who were in prior sex state \( k \).\(^{32}\) Integrating out over the prior sex state leads to the following log likelihood contribution for \( i \)th man of type

\(^{32}\)See Appendix B for how this is calculated.
\[ L_{im}(\theta) = \]
\[ I(y_{im} = 1) \left[ \sum_{w} \sum_{r} I(d_{im} = \{w, r\}) \left( \ln \left[ \sum_{k=0}^{1} \pi_{mk} \phi_{m}^{ur}(\theta, N, k) \right] + \ln \left[ P_{m}^{ur}(\theta, N) \right] \right) \right] \]
\[ + I(y_{im} = 0) \ln \left[ \sum_{k=0}^{1} \sum_{w} \sum_{r} \pi_{mk} \phi_{m}^{ur}(\theta, N, k) \times (1 - P_{m}^{ur}(\theta, N)) \right] \]

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, \( \rho \) and \( A \), are identified through variation in matches across schools with different gender ratios and the overall match rate respectively.\(^{33}\) The estimates of \( \rho \) are significant and negative, ruling out the Cobb-Douglas matching model and confirming that gender ratios do affect the likelihood of observing particular matches. The estimates of \( \rho \) indicate the elasticity of substitution in match production of 0.488.

The middle 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

\(^{33}\)In a 2-market model with only male and female preferences for sex, \( \rho \) and \( A \), the 4 parameters are not identified from only one school. That is because within a school we observe 3 moments: the (1) overall gender ratio, (2) the ratio of men who have sex relative to men who do not have sex (which equals that same ratio for women without sampling error), and (3) the number of men unmatched (only the number of unmatched men or unmatched women is independent since we are counting the gender ratio overall). Thus the search friction \( A \) is identified by including multiple schools.
preferences for sex than females. Those who have had sex in the past also have a much stronger preference for sex in the present.

The lower 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 for age among men. Individuals also prefer to be matched with those in the same grade and minority groups prefer to match with one another. 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. The one surprise is the high value placed on black women, though this value is lower than that placed on black men. The high value for black women occurs because the raw data showed black women matching at a disproportionately high rate given their population size.

5.1 Comparing Model Predictions to Elicited Preferences

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 these model estimated preferences to the stated preferences discussed in Table 2. As the stated preferences were never used in the estimation, a good agreement of the estimated and stated preferences can provide compelling evidence for the credibility of the underlying parameter estimates.

Table 9 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 elicited preferences show 33.2% of women prefer sex to 34.1% of women predicted by the model, while for men the elicited preference for sex is 57.4% compared to a model prediction of 58.1%. The model-predicted preferences for men and women are very close across races for both men and women. The model over-predicts preferences for sex among the ninth graders and under-predicts it among twelfth graders, while it is closer in the intermediate grades.
5.2 Equilibrium Match Probabilities

Table 10 presents the estimated equilibrium match probabilities for whites.\(^\text{34}\) The table is partitioned into 16 cells, one for each possible (female grade, male grade) match. The columns within each cell report the probabilities for men (\(P^m_w\)) and women (\(P^m_w\)). Thus, the upper left cell shows that ninth grade women see significantly higher probabilities of matching than their male classmates. The first column of numbers gives the probabilities that a ninth grade male matches in his eight possible markets (sex or no-sex market crossed with women in four grades). The four columns headed by \(P^m_w\) in the first row of cells give the comparable eight probabilities for ninth grade women. As expected, ninth grade males – in both the sex and no sex markets – see the lowest match probabilities of any group, and these probabilities decline as female grade increases. Note that all 16 cells show women matching with higher probabilities in the sex market and men with higher probabilities in the no sex market. This pattern reflects men’s stronger preference for sex than women’s (Proposition 1.b).

5.3 Counterfactuals

Our structural model permits us to perform counterfactual simulations. We aggregate our sample of 74 Add Health schools to create one large, representative school.\(^\text{35}\) Depending on their grade level, black women in relationships are 7 to 11 percentage points more likely to have sex than their white counterparts. A question we focus on is how much of this gap is driven by differences between the matching markets faced by blacks and whites. Three differences stand out: (i) as compared to whites, amongst blacks the fraction female is much higher and this fraction rises faster with grade level; (ii) the distribution of past sex for black males is far higher than for white males and (iii) the distribution of past sex for black females is higher than for white females. These differences all push the competitive equilibrium toward one in which black women have more sex than their white counterparts.

Our simulated counterfactuals move the matching markets faced by blacks closer to those faced by whites in three cumulative counterfactual steps: (i) first we change the grade-specific gender ratios among blacks to match those of whites, keeping the distribution of other individual characteristics fixed, which is done by removing black women and adding black men while holding the total number of blacks constant; we then (ii) change the distribution of past sex among black men to match that of white men; and finally we also

\(^{34}\)The equilibrium includes individuals from all 32 groups; we present only the subset of probabilities for whites in Table 10

\(^{35}\)The constant returns-to-scale matching technology makes the number of individuals irrelevant.
(iii) change the distribution of past sex among black women to match that of white women.  
Examine the results at each stage tells us to what degree each channel is responsible for the higher rates of sexual participation among blacks.

In Table 11 we present the aggregate school simulated matching probabilities for black men ($P_{wm}^m$) and black women ($P_{wm}^w$) (the simulation includes all individuals of all types, but we present only the probabilities for the sub-set of black-black markets). Within each cell the first two rows contain the baseline probabilities for matching in the sex and no sex markets, while rows three and four report on step (i) (changing the grade-specific gender ratios), and the last two rows report on step (iii) (changing all the conditions of blacks to match those of whites). We omit results from (ii) because they are quite close to (iii).

As seen in the first two rows in the upper left cell, in the baseline aggregate school a ninth grade black female has a dramatically higher probability of matching in the sex market (.687) than her male classmate (.140); and her male classmate can more than double his probability of matching (to .306) by searching in the no sex market. The next two rows (CF Only GR) report on simulation (i) where blacks face the same grade-specific gender ratios as whites. This change increases (decreases) the probabilities of matching in all markets for females (males), but not uniformly. Consistent with Proposition 1, the percentage increase in the probability of matching is higher for women in the market that they prefer, the no sex market. This translates into women changing their search behavior and results in a greater share of matches occurring in the no sex market.

The last two rows “CF-Full” report on (iii) adding to the gender-ratio change an adjustment to the past-sex distributions for both black males and females. This adjustment is a larger change for black males than for black females as the racial gap in past sex is larger for males. It results in a decreased number of sexually experienced black men, reducing the competition among men in the sex-market, and increasing their probability of matching in the sex market. Correspondingly, the probability of men matching in the no-sex market decreases. The opposite changes occur for women: as compared to the case where just the gender ratio was changed, women see higher probabilities of matching in the no sex market and lower probabilities of matching in the sex market. With fewer black men having had sex in the past, demand decreases in the sex market, lowering female probabilities of matching there and correspondingly increasing female probabilities of matching in the no sex market.

---

36 First changing the female past-sex distribution, and then the male past-sex distribution generated the same pattern of results: the male past-sex distribution among blacks is the major driver of the higher sexual participation.

37 The table shows only increases for blacks, other races will qualitatively see the same pattern.

38 For example in the black, 9th-grade market the percentage change in the no-sex market was .595% and the change in the sex market was .384%. The magnitude was greater in the no-sex markets for partner-characteristics.
Thus (iii) magnifies the change in women’s behavior seen in (i) and the result is an even greater share of matches occurring in the no sex market.

Table 12 captures how these changes in the probabilities of matching translate into changes in the fraction of relationships with sexual engagement. For expositional reasons we report a racial gap in sexual engagement for same-grade and race matches only. The “racial gap” compares probabilities of a match in the sex market conditional on matching. Table 12 reports the difference between these conditional probabilities for whites matched with whites minus that for blacks matched with blacks. The baseline shows the 7 to 9 percentage point racial gap noted earlier. When we simulate (i) blacks facing the same grade-specific gender-ratios as whites, the gap shrinks, particularly for younger couples, but the effects are small. Additionally adjusting the past-sex distribution for black men (ii) yields a much larger effect, up to 75% for 12th graders and over 50% for all groups. Finally when we additionally adjust the distribution of past sex among black females to match that among white females (iii), the remaining racial gap is extremely small and is explained by differences in preference over same-race and grade. These full simulations show that a remarkable 55% to 75% of the racial gap in sexual participation between black and white women can be attributed to the differences in the matching markets they face.

5.4 Robustness Checks

We now 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:

\[
\Pr(w, r|m) = \frac{\exp \left( \frac{\mu_{m}^{w}r + \ln(P_{m}^{w})}{\sigma} \right)}{\sum_{w'}\sum_{r'} \exp \left( \frac{\mu_{m}'r' + \ln(P_{m}^{w'})}{\sigma} \right) + \exp \left( \frac{\mu_{0}^{m}}{\sigma} \right)}
\]

\( \text{Equation 13} \)

\( ^{39} \)Results look similar within a partner grade category.
Using the independence of irrelevant alternatives (IIA) property, note this probability conditional on searching in the school is:

\[
\Pr(w, r|m, d \neq 0) = \phi_m^{\text{wt}} = \frac{\exp\left(\mu_m^{\text{wt}} + \ln(P_m^{\text{wt}})\right)}{\sum_{w'} \sum_{r'} \exp\left(\mu_{m'}^{w'} r' + \ln(P_m^{w'} r')\right)},
\]

which is the probability in equation (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 \( N_0^m \) 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 \( N_0^m \). With the probability of matching in the outside market given by \( P_0^m \), instead of forming \( N_m \) by subtracting off \( N_0^m \) for the population, we subtract off \( N_0^m / P_0^m \). 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.\(^{40}\)

Results for the estimated model under different match rate assumptions are presented in Table 13, with the first column repeating the results of the original model. The top rows of Table 13 show the mean choice probability for choosing sex without equilibrium influences.

\(^{40}\)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 \( F_0^m \), \( F_0^w \), we inflate this fraction by one over match rate outside the 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 \( N_0^m \), but estimate it as \( N_0^m = F_0^m N_m \).
For reasonable match rates, we continue to see the same patterns of women preferring sex relative to men. However, as the match rate falls, \( \rho \) moves closer to zero, which is the Cobb-Douglas case.\(^{41}\)

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 to 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 13. 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 prior to Wave I, 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 13). The results are presented in the second and third sets of columns 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. It is only in going to the extreme (the final column) of having individuals face a 50% probability of matching in the outside market, and when we assume very few individuals are searching for a first-time relationship, do we begin to lower the estimated preferences for men and raise them for women. Even with a fairly large friction in the outside market the predicted preferences do not fit the subjectives very closely. There may be some optimal combination of adjustment along these two dimensions which yields model fitness similar to the baseline, but over all the specifications the qualitative pattern of results (e.g. sex preferences, age profiles and cross-race matching) look best under our baseline assumptions that everyone searches and that matching outside the school is close the frictionless.

\(^{41}\)We estimated models with lower out-matching fractions, but \( \rho \) becomes very small and insignificant erasing the identifying power of the gender ratios for uncovering gender differences in utility.
6 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 relationship terms using variation in the gender ratio. When the researcher’s goal is to understand differences in male and female preferences, directed search provides an attractive 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 matching market and uncovered male and female differences in preferences for sex. The preferences from the structural model match the self-reported preferences, providing a compelling out-of-sample test for the validity for the approach. That men and women value sex differently suggests that changes in sexual behaviors may have different welfare effects for men than for women. Further, when gender ratios tilt such that men become a minority—as, for example, 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. For high school students our counterfactual simulations show that, conditional on matching, most of the gap in sexual engagement between black and white women is driven by the unfavorable market conditions that black women face. If conditions faced by blacks (as measured by the gender ratio and sexual experience of males) were similar to those for whites, the racial gap in sexual participation would shrink a remarkable 50 to 75 percent.

Beyond Teen Sex

The result that individuals trade off preferred characteristics of a match for an increased chance of matching touches areas of concern well beyond matching in U.S. high schools. To indicate the scope for employing a model such as ours, we enumerate some of these areas.

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, given 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 for raising them. That women faced with a shortage of partners respond by demanding less favorable terms in their relationships (in this case, fathers bearing little responsibility for child-rearing) is consistent with our approach.

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 for her recent cohorts of highly educated women, Rose (2004) found that 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, following the carnage of World War I, France’s shortage of men resulted in surviving French men marrying into a higher social classes after the war than comparable French men did prior to 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. Typically, western countries have about 1,050 women per 1,000 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 thousand men. Within Punjab the district of Ludhaina had only 824 women per thousand men. Such endogenously determined severe imbalances will no doubt greatly impact marriage patterns, dowries, intra-household resource distributions, and fertility, thereby differentially impacting the well-being of men and women.

Finally, 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 indicate that women’s wellbeing has declined both relative to men’s as well as absolutely. They found this across a variety of data sources, subjective measures of well-being, demographic groups and industrialized countries. Studies such as ours can begin to explain such apparent paradoxes.

References


42 For example, following the famous Sen (1990) 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.


Appendix A

Proof of Proposition 1 The proof of claim (a) follows from manipulating the definition of the search probability. Assuming $\mu_{mr}' - \mu_{mr} > \mu_{wr}' - \mu_{wr}$, we can add the log-match probability for each combination to both sides in the following way:

$$\mu_{mr}' + \log(P_{mr}') - \mu_{mr} - \log(P_{mr}) + \log(P_{wr}') - \log(P_{wr}) >$$

$$\mu_{wr}' + \log(P_{wr}') - \mu_{wr} - \log(P_{mr}') + \log(P_{w}) - \log(P_{mr}).$$

(15)

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

$$\frac{e^{\mu_{mr}' + \log(P_{mr}')}}{e^{\mu_{mr} + \log(P_{mr})}} e^{\log(P_{mr}') - \log(P_{mr})} >$$

$$\frac{e^{\mu_{wr}' + \log(P_{wr}')}}{e^{\mu_{wr} + \log(P_{mr})}} e^{\log(P_{mr}') - \log(P_{mr})}$$

(16)
or:

\[
\frac{\phi_{mr}' P_{mr}'}{\phi_{mr} P_{mr}} > \frac{\phi_{mr}' P_{mr}}{\phi_{mr} P_{mr}'}.
\]

(17)

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

\[
\frac{P_{mr}'}{P_{mr}} = \frac{\left[ 1 + \left( \frac{\phi_{mr}' N_w}{\phi_{mr} N_m} \right) \rho \right]^{1/\rho}}{\left[ 1 + \left( \frac{\phi_{mr}' N_m}{\phi_{mr} N_w} \right) \rho \right]^{1/\rho}}
\]

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

\[
\frac{P_{mr}'}{P_{mr}} = \left[ \frac{\phi_{mr}' N_w}{\phi_{mr} N_m} \right]^{1/\rho}.
\]

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

\[
\frac{\phi_{mr}'}{\phi_{mr}} \left[ \frac{\phi_{mr}' N_w}{\phi_{mr} N_m} \right]^{1/\rho} > \frac{\phi_{mr}'}{\phi_{mr}} \left[ \frac{\phi_{mr}' N_w}{\phi_{mr} N_m} \right]^{1/\rho},
\]

(17)

which further simplifies to:

\[
\left( \frac{\phi_{mr}'}{\phi_{mr}} \right)^2 > \left( \frac{\phi_{mr}'}{\phi_{mr}} \right)^2,
\]

(17)

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:

\[
\frac{\phi_{mr}}{\phi_{mr}'} < \frac{\phi_{mr}'}{\phi_{mr}}
\]

(17)

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

\[
\left( \frac{\phi_{mr} N_w}{\phi_{mr}' N_m} \right)^{1/\rho} > \left( \frac{\phi_{mr}' N_w}{\phi_{mr} N_m} \right)^{1/\rho}
\]

(17)

adding one to both sides and raising both to the power \(\rho\) switches the inequality once more and we have:

\[
\left[ 1 + \left( \frac{\phi_{mr} N_w}{\phi_{mr}' N_m} \right)^{1/\rho} \right]^{\rho} < \left[ 1 + \left( \frac{\phi_{mr}' N_w}{\phi_{mr} N_m} \right)^{1/\rho} \right]^{\rho}
\]

(17)

which is the definition of \(P_{mr} < P_{mr}'\). 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:

\[
\ln(\phi_{wr}^{mr}) - \ln(1-\phi_{wr}^{mr}) \equiv \ln(\mu_{wr}^{mr}) - \ln(\mu_{wr}^{mr}) + \ln \left[ 1 + \left( \frac{\phi_{wr}^{mr} G}{\phi_{wr}^{mr}} \right)^p \right] - \ln \left[ 1 + \left( \frac{(1 - \phi_{wr}^{mr}) G}{1 - \phi_{wr}^{mr}} \right)^p \right]
\]

(17)

Now, define \( F_1(\phi_{wr}^{mr}, \phi_{wr}^{mr}, G) \) and \( F_2(\phi_{wr}^{mr}, \phi_{wr}^{mr}, G) \) based on the identity in 18 and the corresponding expression for men, respectively:

\[
F_1(\phi_{wr}^{mr}, \phi_{wm}^{mr}, G) \equiv \ln(\phi_{wr}^{mr}) - \ln(1-\phi_{wr}^{mr}) - \ln(\mu_{wr}^{mr}) + \ln(\mu_{wr}^{mr}) - \ln \left[ 1 + \left( \frac{\phi_{wr}^{mr} G}{\phi_{wm}^{mr}} \right)^p \right] + \ln \left[ 1 + \left( \frac{(1 - \phi_{wm}^{mr}) G}{1 - \phi_{wm}^{mr}} \right)^p \right]
\]

(18)

\[
F_2(\phi_{wr}^{mr}, \phi_{wm}^{mr}, G) \equiv \ln(\phi_{wm}^{mr}) - \ln(1-\phi_{wm}^{mr}) - \ln(\mu_{wm}^{mr}) + \ln(\mu_{wm}^{mr}) - \ln \left[ 1 + \left( \frac{\phi_{wm}^{mr} G}{\phi_{wm}^{mr}} \right)^p \right] + \ln \left[ 1 + \left( \frac{(1 - \phi_{wm}^{mr}) G}{1 - \phi_{wm}^{mr}} \right)^p \right]
\]

(18)

which can equivalently be expressed as:

\[
F_1(\phi_{wr}^{mr}, \phi_{wm}^{mr}, G) \equiv 2 \ln(\phi_{wr}^{mr}) - 2 \ln(1-\phi_{wr}^{mr}) - \ln(\mu_{wr}^{mr}) + \ln(\mu_{wr}^{mr}) - \ln \left[ (\phi_{wr}^{mr})^p + (\phi_{wm}^{mr} G)^p \right] + \ln \left[ (1 - \phi_{wm}^{mr})^p + (1 - \phi_{wm}^{mr} G)^p \right]
\]

(18)

\[
F_2(\phi_{wr}^{mr}, \phi_{wm}^{mr}, G) \equiv 2 \ln(\phi_{wm}^{mr}) - 2 \ln(1-\phi_{wm}^{mr}) - \ln(\mu_{wm}^{mr}) + \ln(\mu_{wm}^{mr}) - \ln \left[ (\phi_{wm}^{mr})^p + (\phi_{wm}^{mr} G)^p \right] + \ln \left[ (1 - \phi_{wm}^{mr})^p + (1 - \phi_{wm}^{mr} G)^p \right]
\]

(18)

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

\[
\begin{bmatrix}
\frac{\partial F_1}{\partial \phi_w^{mr}} & \frac{\partial F_1}{\partial \phi_w^{mr}} \\
\frac{\partial F_2}{\partial \phi_w^{mr}} & \frac{\partial F_2}{\partial \phi_w^{mr}}
\end{bmatrix}
= - \begin{bmatrix}
\frac{\partial F_1}{\partial G} \\
\frac{\partial F_2}{\partial G}
\end{bmatrix}
\]

which can be expressed as:

\[
\begin{bmatrix}
\frac{\partial \phi_w^{mr}}{\partial G} & \frac{\partial \phi_w^{mr}}{\partial G} \\
\frac{\partial \phi_w^{mr}}{\partial G} & \frac{\partial \phi_w^{mr}}{\partial G}
\end{bmatrix}
= - \left( \begin{bmatrix}
\frac{\partial F_1}{\partial \phi_w^{mr}} & \frac{\partial F_1}{\partial \phi_w^{mr}} \\
\frac{\partial F_2}{\partial \phi_w^{mr}} & \frac{\partial F_2}{\partial \phi_w^{mr}}
\end{bmatrix} \right)^{-1}
\begin{bmatrix}
\frac{\partial F_1}{\partial G} \\
\frac{\partial F_2}{\partial G}
\end{bmatrix}
\]

34
The partial derivatives of $F_1$ and $F_2$ with respect to $G$ can be expressed as:

\[
\frac{\partial F_1}{\partial G} = G^{-1} \left( \left[ \left( \frac{1 - \phi_{mr}^*}{1 - \phi_{mr}^*} \right)^\rho + 1 \right]^{-1} - \left[ \left( \frac{\phi_{mr}^*}{\phi_{mr}^*} \right)^\rho + 1 \right]^{-1} \right)
\]

(18)

\[
\frac{\partial F_2}{\partial G} = G^{-1} \left( \left[ \left( \frac{\phi_{mr}^*}{\phi_{mr}^*} \right)^\rho + 1 \right]^{-1} - \left[ \left( 1 - \frac{\phi_{mr}^*}{\phi_{mr}^*} \right) \right]^{-1} \right)
\]

(19)

Since $\rho < 0$ and (by claim (a)) $\frac{1 - \phi_{mr}^*}{1 - \phi_{mr}^*} = \frac{\phi_{mr}^*}{\phi_{mr}^*} < \frac{\phi_{mr}^*}{\phi_{mr}^*}$, 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 $\phi_{mr}^*$ and $\phi_{mr}^*$ can be expressed as:

\[
\frac{\partial F_1}{\partial \phi_{mr}^*} = \frac{(\phi_{mr}^*)^\rho + 2(\phi_{mr}^* G)^\rho}{(\phi_{mr}^*)^\rho + (\phi_{mr}^* G)^\rho} + \frac{(1 - \phi_{mr}^*)^\rho + 2((1 - \phi_{mr}^*) G)^\rho}{(1 - \phi_{mr}^*)^\rho + ((1 - \phi_{mr}^*) G)^\rho} > 0
\]

(21)

\[
\frac{\partial F_1}{\partial \phi_{mr}^*} = -\frac{(\phi_{mr}^*)^\rho - (\phi_{mr}^* G)^\rho}{(\phi_{mr}^*)^\rho + (\phi_{mr}^* G)^\rho} + \frac{(1 - \phi_{mr}^*)^\rho - ((1 - \phi_{mr}^*) G)^\rho}{(1 - \phi_{mr}^*)^\rho + ((1 - \phi_{mr}^*) G)^\rho} < 0
\]

(22)

\[
\frac{\partial F_2}{\partial \phi_{mr}^*} = \frac{2(\phi_{mr}^*)^\rho + (\phi_{mr}^* G)^\rho}{(\phi_{mr}^*)^\rho + (\phi_{mr}^* G)^\rho} + \frac{2((1 - \phi_{mr}^*)^\rho + ((1 - \phi_{mr}^*) G)^\rho}{(1 - \phi_{mr}^*)^\rho + ((1 - \phi_{mr}^*) G)^\rho} > 0
\]

(23)

\[
\frac{\partial F_2}{\partial \phi_{mr}^*} = -\frac{2(\phi_{mr}^*)^\rho - (\phi_{mr}^* G)^\rho}{(\phi_{mr}^*)^\rho + (\phi_{mr}^* G)^\rho} + \frac{2((1 - \phi_{mr}^*)^\rho - ((1 - \phi_{mr}^*) G)^\rho}{(1 - \phi_{mr}^*)^\rho + ((1 - \phi_{mr}^*) G)^\rho} < 0
\]

(24)

Appealing to Cramer’s rule,

\[
\frac{\partial \phi_{mr}^*}{\partial G} = \frac{\partial F_2}{\partial \phi_{mr}^*} \frac{\partial G}{\partial \phi_{mr}^*} - \frac{\partial F_2}{\partial \phi_{mr}^*} \frac{\partial \phi_{mr}^*}{\partial G}
\]

(25)

\[
\frac{\partial \phi_{mr}^*}{\partial G} = \frac{\partial F_1}{\partial \phi_{mr}^*} \frac{\partial G}{\partial \phi_{mr}^*} - \frac{\partial F_1}{\partial \phi_{mr}^*} \frac{\partial \phi_{mr}^*}{\partial G}
\]

(26)

In both cases, the numerators are positive. Both have one negative term, $\partial F_1/\partial \phi_{mr}^*$ and $\partial F_2/\partial \phi_{mr}^*$ 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.
Appendix B

In this appendix we discuss three issues with the data and one issue with the scaling of the error term. The three issues with the data are i) determining the share of students searching in the outside market, ii) determining the distribution of prior sex for males, and iii) cases where females do not report characteristics of their partners. The scaling of error term comes about because the Type 1 extreme value shocks naturally lead to minority groups being more sought after above and beyond the gender ratio.

To deal with the fraction searching outside the school we begin with a strong assumption and subsequently relax it. We assume initially that each individual could match outside the school with probability one. This means that we only need the fraction of each individual type matching outside the school to correct the aggregate gender ratios to reflect the number of men and women of each type searching in the school. We estimate (separately for men and women) a logit on matching outside the school which is a function of individual grade, race and school fixed effects. So for instance for men we specify the probability of matching outside the school for an \(m\) type man at school \(s\) as:

\[
P(\text{MatchOut}|m,s) = \frac{\exp \left( \sum_g I(G_m = g) \gamma_g + \sum_r I(R_m = r) \gamma_r + \gamma_s \right)}{1 + \exp \left( \sum_g I(G_m = g) \gamma_g + \sum_r I(R_m = r) \gamma_r + \gamma_s \right)}.
\]

(26)

The resulting predicted conditional probabilities are used to scale down the number of searching men \(m\) and searching women \(w\) within in each school. We subsequently relax the perfect ability to match outside the school by imposing that each match observed required more individuals searching outside the school in order to materialize (e.g. if we see that one male of type \(m\) matched outside the school, assuming the probability of matching outside was one-half, we would reduce the number of type \(m\) men searching in the school by two).

To deal with the unobserved distribution of male past sexual activity from using female reports, we estimate the conditional probability of past sex at the school level from the male half of the original sample. We do this only among those who are not matched outside the school, thus we specify the probability of past sex for an \(m\) type man at school \(s\) as:

\[
P(\text{PastSex}|m,s,\text{MatchOut} = 0) = \frac{\exp \left( \sum_g I(G_m = g) \theta_g + \sum_r I(R_m = r) \theta_r + \theta_s \right)}{1 + \exp \left( \sum_g I(G_m = g) \theta_g + \sum_r I(R_m = r) \theta_r + \theta_s \right)} = \frac{\pi_{m1}^s}{\pi_{m1}}.
\]

(26)

again using grade, race and school fixed effects. These predicted conditional probabilities are used as weights to integrate out the likelihood contribution for men in equation (12).
The final data issue concerns missing reports on partner characteristics. Because we use the female-reported match distribution we are unable to recreate matches when females fail to report either their partners’ grade or race. We remain agnostic about how these individuals enter the matching problem and proceed after dropping them from our sample, thus our results only rationalize the match distribution conditional on full reporting.\footnote{Female censoring was slightly lower than male censoring. Imputing censored match characteristics from the observed distribution within or across schools introduced attenuation in the relationship between gender ratios and matching, suggesting the data contain too little information to identify the parameters of interest (in particular $\rho$) when including censored observations.} Results were quite similar when including them as unmatched or integrating out over the male characteristics of female matches.\footnote{This last approach requires a less restrictive assumption: that male observations lost to censoring are randomly distributed with respect to male characteristics and terms.}

The one remaining issue is then scaling the unobserved preferences. The baseline model is set up so that individuals get one draw on their unobserved preference for each race-grade-relationship combination, regardless of how many potential partners there are. Hence, a black woman moving to a predominately white school will see her match probabilities rise because she is in such a minority relative to white women. However, this is not what we see in the data. For example, estimating a logit model on the probability of being in a relationship for blacks does not show a positive relationship with either fraction white or fraction non-black.

To deal with this issue, we add a term to the utility function for a particular market which is the log share of the number of the school population in that race-grade combination. Hence, in the market for 9th grade white women, this adjustment term is the log of the share of the women at the school who are white and in 9th grade. This adjustment term is the same for the sex and no sex market. Adjusting the utility in this way is equivalent to a model where the number of draws on the unobserved preferences for each market type is weighted by the proportion of that type in the school, meaning the logistic assumption is much less likely to drive individuals to search in markets with very few individuals.\footnote{Kennan and Walker (2011) faced a similar issue in examining migration decisions. A similar adjustment was made such that states with large populations emitted more draws for the unobserved preference for their state.}
Table 1: Means by Gender$^a$

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currently</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matched (sex or relationship)</td>
<td>0.311</td>
<td>0.296</td>
</tr>
<tr>
<td>In a Relationship</td>
<td>0.302</td>
<td>0.288</td>
</tr>
<tr>
<td>Having Sex</td>
<td>0.160</td>
<td>0.142</td>
</tr>
<tr>
<td>Prior sex</td>
<td>0.306</td>
<td>0.220</td>
</tr>
<tr>
<td>Current Sex</td>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.167</td>
<td>0.155</td>
</tr>
<tr>
<td>Black</td>
<td>0.200</td>
<td>0.171</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.147</td>
<td>0.090</td>
</tr>
<tr>
<td>Other</td>
<td>0.083</td>
<td>0.073</td>
</tr>
<tr>
<td>Prior Sex</td>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.261</td>
<td>0.218</td>
</tr>
<tr>
<td>Black</td>
<td>0.500</td>
<td>0.300</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.354</td>
<td>0.174</td>
</tr>
<tr>
<td>Other</td>
<td>0.186</td>
<td>0.119</td>
</tr>
<tr>
<td>N</td>
<td>3,747</td>
<td>3,533</td>
</tr>
</tbody>
</table>

$^a$Sample includes only those searching in-school, under the assumption that $p_{match} = 1$. Current is defined as ongoing at the time of the in-home survey. Relationship means holding hands and kissing.
Table 2: Stated Preferences by Gender<sup>a</sup>

<table>
<thead>
<tr>
<th>Prefer:</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationship</td>
<td>0.907</td>
<td>0.916</td>
</tr>
<tr>
<td>Sex</td>
<td>0.574</td>
<td>0.341</td>
</tr>
</tbody>
</table>

Sex, By Race:

<table>
<thead>
<tr>
<th>Race</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>0.557</td>
<td>0.334</td>
</tr>
<tr>
<td>Black</td>
<td>0.690</td>
<td>0.377</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.612</td>
<td>0.354</td>
</tr>
</tbody>
</table>

Sex, By Grade:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>9th</td>
<td>0.456</td>
<td>0.237</td>
</tr>
<tr>
<td>10th</td>
<td>0.554</td>
<td>0.328</td>
</tr>
<tr>
<td>11th</td>
<td>0.648</td>
<td>0.369</td>
</tr>
<tr>
<td>12th</td>
<td>0.685</td>
<td>0.483</td>
</tr>
</tbody>
</table>

N 3,747 3,533

<sup>a</sup>Answers come from questions on whether the respondents’ ideal relationship over the coming year would include sex or a relationship as defined above.

Table 3: Cross-Grade Matching Distribution.<sup>a</sup>

<table>
<thead>
<tr>
<th>Female Grade</th>
<th>Male Grade</th>
<th>Total</th>
<th>Fraction of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9th</td>
<td>10th</td>
<td>11th</td>
</tr>
<tr>
<td>9th</td>
<td>0.085</td>
<td>0.066</td>
<td>0.057</td>
</tr>
<tr>
<td>10th</td>
<td>0.029</td>
<td>0.096</td>
<td>0.085</td>
</tr>
<tr>
<td>11th</td>
<td>0.020</td>
<td>0.040</td>
<td>0.095</td>
</tr>
<tr>
<td>12th</td>
<td>0.005</td>
<td>0.017</td>
<td>0.065</td>
</tr>
</tbody>
</table>

Total 0.138 0.219 0.302 0.341 1.000

<sup>a</sup>Distribution from 1047 within-school matches. Fraction of sample refers to sample of in-school searchers under the assumption that $P_{\text{out match}} = 1$. 
Table 4: Cross-Race Matching Distribution.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Female Race</th>
<th>White</th>
<th>Male Race</th>
<th>Black</th>
<th>Hispanic</th>
<th>Other</th>
<th>Total</th>
<th>Fraction of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>0.572</td>
<td>Black</td>
<td>0.010</td>
<td>Hispanic</td>
<td>0.041</td>
<td>Other</td>
<td>0.009</td>
</tr>
<tr>
<td>Black</td>
<td>0.008</td>
<td>Hispanic</td>
<td>0.183</td>
<td>0.010</td>
<td>0.009</td>
<td>Other</td>
<td>0.000</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.031</td>
<td>Other</td>
<td>0.008</td>
<td>Hispanic</td>
<td>0.081</td>
<td>Other</td>
<td>0.001</td>
</tr>
<tr>
<td>Other</td>
<td>0.011</td>
<td>Total</td>
<td>0.007</td>
<td>0.010</td>
<td>0.020</td>
<td>Other</td>
<td>0.048</td>
</tr>
<tr>
<td>Total</td>
<td>0.622</td>
<td></td>
<td>0.207</td>
<td>0.141</td>
<td>0.030</td>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td>Fraction of Sample</td>
<td>0.600</td>
<td></td>
<td>0.162</td>
<td>0.158</td>
<td>0.081</td>
<td></td>
<td>1.000</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Distribution from 1047 within-school matches. Fraction of sample refers to sample of in-school searchers under the assumption that $P_{\text{match}} = 1$.

Table 5: Conditional Means of Sex Participation\textsuperscript{a}

<table>
<thead>
<tr>
<th>Observed:</th>
<th>Women</th>
<th>Men</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(Sex</td>
<td>Want sex=1)</td>
<td>0.304</td>
<td>0.252</td>
</tr>
<tr>
<td>P(Sex</td>
<td>Want sex=1,Matched)</td>
<td>0.716</td>
<td>0.649</td>
</tr>
<tr>
<td>P(Sex</td>
<td>Want sex=0,Matched)</td>
<td>0.262</td>
<td>0.196</td>
</tr>
</tbody>
</table>

With No Prior Sex

<table>
<thead>
<tr>
<th>Observed:</th>
<th>Women</th>
<th>Men</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(Sex</td>
<td>Want sex=1,Matched)</td>
<td>0.579</td>
<td>0.433</td>
</tr>
<tr>
<td>P(Sex</td>
<td>Want sex=0,Matched)</td>
<td>0.177</td>
<td>0.123</td>
</tr>
</tbody>
</table>

With Prior Sex

<table>
<thead>
<tr>
<th>Observed:</th>
<th>Women</th>
<th>Men</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(Sex</td>
<td>Want sex=1,Matched)</td>
<td>0.884</td>
<td>0.850</td>
</tr>
<tr>
<td>P(Sex</td>
<td>Want sex=0,Matched)</td>
<td>0.709</td>
<td>0.516</td>
</tr>
</tbody>
</table>

\textsuperscript{a}*,**,*** denote significance at the 5,1, and 0.01% levels respectively. Matched is defined as having either a relationship or sex in-school. Sample includes only in-school searchers under the assumption that $P_{\text{match}} = 1$. 

40
Table 6: Variation in Fraction Female\textsuperscript{a}

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

Overall Fraction Female

| P(Sex|Match) | < Median | > Above Median |
|-----------|----------|----------------|
| N         | 555      | 492            |

Same-Race Fraction Female

| P(Sex|Match) | < Median | > Above Median |
|-----------|----------|----------------|
| N         | 544      | 503            |

Fraction Female of Partner’s Race-Grade

| P(Sex|Match) | < Median | > Above Median |
|-----------|----------|----------------|
| N         | 534      | 513            |

\textsuperscript{a}Based on a sample of 74 schools. Gender ratios are calculated using only those searching within the school. Probability of sex conditional on matching is calculated from only in-school matches. Aggregate gender ratio refers to the fraction of the searching population that is female.
Table 7: Reduced Form Probability of Sex Conditional on Matching

|                               | P(Sex|Match)       |       |       |
|-------------------------------|---------------|-------|-------|
|                               | (i)           | (ii)  | (iii) |
| Fraction Female of Partner’s Race-Grade | 0.183*        | 0.266** | 0.313** |
|                               | (0.119)       | (0.118) | (0.138) |
| Prior Sex                     | 0.514***      | 0.493*** | 0.495*** |
|                               | (0.094)       | (0.092) | (0.097) |
| Grade                         | 0.082***      | 0.084*** | 0.081*** |
|                               | (0.013)       | (0.012) | (0.013) |
| Partner Grade                 | 0.038**       | 0.040** | 0.038** |
|                               | (0.016)       | (0.016) | (0.017) |
| School Characteristics        | No            | Yes   | No    |
| School Fixed Effects          | No            | No    | Yes   |
| N                             | 1046          | 1046  | 1039  |

*aCoefficients are probit marginal effects from the probability of having sex conditional on matching. Regressions are run for females, and include only in-school searchers. All specifications include linear grade and partner grade, prior sex, prior sex interacted with own and partner grade, and indicators for each own and partner race-combination. School characteristics are: are percent non-white, total males and females with no prior sex and with prior sex. *,**,*** denote significance at the 10, 5, and 1% levels respectively.
Table 8: Structural Model Estimates

<table>
<thead>
<tr>
<th>Matching Parameters</th>
<th>Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>-1.051</td>
<td>(0.222)</td>
</tr>
<tr>
<td>A</td>
<td>0.418</td>
<td>(0.012)</td>
</tr>
<tr>
<td>$\sigma^{-1}$</td>
<td>0.355</td>
<td>(0.074)</td>
</tr>
</tbody>
</table>

Sex Utility

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male $\times$ Sex ($\alpha_7$)</td>
<td>-0.688</td>
<td>(0.461)</td>
</tr>
<tr>
<td>Female $\times$ Sex ($\alpha_9$)</td>
<td>-3.766</td>
<td>(0.698)</td>
</tr>
<tr>
<td>Past-Sex $\times$ Sex ($\alpha_8$)</td>
<td>7.316</td>
<td>(1.138)</td>
</tr>
</tbody>
</table>

Match Utility

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same grade ($\alpha_1$)</td>
<td>1.911</td>
<td>(0.320)</td>
</tr>
<tr>
<td>Partner Grade $\times$ Boy ($\alpha_2$)</td>
<td>0.515</td>
<td>(0.130)</td>
</tr>
<tr>
<td>Partner Grade $\times$ Girl ($\alpha_5$)</td>
<td>2.629</td>
<td>(0.328)</td>
</tr>
<tr>
<td>Same Race ($\alpha_3$)</td>
<td>4.342</td>
<td>(0.735)</td>
</tr>
<tr>
<td>Partner Black $\times$ Boy ($\alpha_{4b}$)</td>
<td>4.094</td>
<td>(0.601)</td>
</tr>
<tr>
<td>Partner Black $\times$ Girl ($\alpha_{6b}$)</td>
<td>5.302</td>
<td>(0.627)</td>
</tr>
<tr>
<td>Partner Hisp $\times$ Boy ($\alpha_{4h}$)</td>
<td>-1.998</td>
<td>(0.714)</td>
</tr>
<tr>
<td>Partner Hisp $\times$ Girl ($\alpha_{6h}$)</td>
<td>-0.429</td>
<td>(0.589)</td>
</tr>
<tr>
<td>Partner Other $\times$ Boy ($\alpha_{4o}$)</td>
<td>-3.877</td>
<td>(0.910)</td>
</tr>
<tr>
<td>Partner Other $\times$ Girl ($\alpha_{6o}$)</td>
<td>-6.682</td>
<td>(1.197)</td>
</tr>
</tbody>
</table>

*Estimates are from sample of in-school searchers 7355 individuals and 1047 two-sided matches. Standard errors are in parentheses.
Table 9: Stated vs. Predicted Preferences for Sex

<table>
<thead>
<tr>
<th>Group:</th>
<th>Stated</th>
<th>Model Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.574</td>
<td>0.581</td>
</tr>
<tr>
<td>Female</td>
<td>0.341</td>
<td>0.332</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Male</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>0.557</td>
<td>0.561</td>
</tr>
<tr>
<td>Black</td>
<td>0.690</td>
<td>0.668</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.612</td>
<td>0.600</td>
</tr>
<tr>
<td>9th Grade</td>
<td>0.456</td>
<td>0.545</td>
</tr>
<tr>
<td>10th Grade</td>
<td>0.554</td>
<td>0.572</td>
</tr>
<tr>
<td>11th Grade</td>
<td>0.648</td>
<td>0.603</td>
</tr>
<tr>
<td>12th Grade</td>
<td>0.685</td>
<td>0.620</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Female</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>0.334</td>
<td>0.331</td>
</tr>
<tr>
<td>Black</td>
<td>0.377</td>
<td>0.377</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.354</td>
<td>0.304</td>
</tr>
<tr>
<td>9th Grade</td>
<td>0.237</td>
<td>0.281</td>
</tr>
<tr>
<td>10th Grade</td>
<td>0.328</td>
<td>0.327</td>
</tr>
<tr>
<td>11th Grade</td>
<td>0.369</td>
<td>0.357</td>
</tr>
<tr>
<td>12th Grade</td>
<td>0.483</td>
<td>0.387</td>
</tr>
</tbody>
</table>

Subjective preference means come from sample of both men (N=3689) and women (N=3472). The model predicted means set the probability of matching to one, giving the average choice probability across individuals based only on preferences absent matching concerns.
Table 10: Equilibrium Probabilities of Matching: Whites$^a$

<table>
<thead>
<tr>
<th>Female Grade and Relationship</th>
<th>Male Grade</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>9 sex</td>
<td>$P^w_m$</td>
<td>$P^m_w$</td>
<td>$P^w_m$</td>
<td>$P^m_w$</td>
<td>$P^w_m$</td>
<td>$P^m_w$</td>
<td>$P^w_m$</td>
</tr>
<tr>
<td>9 no sex</td>
<td>0.109</td>
<td>0.715</td>
<td>0.228</td>
<td>0.604</td>
<td>0.402</td>
<td>0.435</td>
<td>0.573</td>
</tr>
<tr>
<td>10 sex</td>
<td>0.094</td>
<td>0.728</td>
<td>0.201</td>
<td>0.629</td>
<td>0.367</td>
<td>0.470</td>
<td>0.543</td>
</tr>
<tr>
<td>10 no sex</td>
<td>0.164</td>
<td>0.664</td>
<td>0.332</td>
<td>0.503</td>
<td>0.535</td>
<td>0.300</td>
<td>0.678</td>
</tr>
<tr>
<td>11 sex</td>
<td>0.068</td>
<td>0.752</td>
<td>0.152</td>
<td>0.676</td>
<td>0.295</td>
<td>0.539</td>
<td>0.473</td>
</tr>
<tr>
<td>11 no sex</td>
<td>0.111</td>
<td>0.713</td>
<td>0.245</td>
<td>0.588</td>
<td>0.443</td>
<td>0.394</td>
<td>0.616</td>
</tr>
<tr>
<td>12 sex</td>
<td>0.047</td>
<td>0.770</td>
<td>0.109</td>
<td>0.715</td>
<td>0.225</td>
<td>0.607</td>
<td>0.392</td>
</tr>
<tr>
<td>12 no sex</td>
<td>0.073</td>
<td>0.747</td>
<td>0.171</td>
<td>0.658</td>
<td>0.345</td>
<td>0.491</td>
<td>0.535</td>
</tr>
</tbody>
</table>

$^a$Each cell gives the probability of matching in sex or no sex markets based on an individuals’ grade and possible partner grade. $P^w_m$ is the probability of matching for a man looking for a woman, $P^m_w$ is the probability of matching for a woman looking for a man.
Table 11: Counterfactual Probabilities of Matching: Blacks in Aggregate School

<table>
<thead>
<tr>
<th>Female Grade and Relationship:</th>
<th>Male Grade</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P_w^m$</td>
<td>$P_w^m$</td>
<td>$P_m^w$</td>
<td>$P_m^w$</td>
<td>$P_w^m$</td>
<td>$P_w^m$</td>
<td>$P_m^w$</td>
<td>$P_m^w$</td>
<td>$P_w^m$</td>
<td>$P_w^m$</td>
<td>$P_m^w$</td>
<td>$P_m^w$</td>
</tr>
<tr>
<td>sex baseline</td>
<td>0.140</td>
<td>0.687</td>
<td>0.285</td>
<td>0.550</td>
<td>0.465</td>
<td>0.372</td>
<td>0.619</td>
<td>0.212</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex baseline</td>
<td>0.306</td>
<td>0.529</td>
<td>0.519</td>
<td>0.316</td>
<td>0.677</td>
<td>0.151</td>
<td>0.753</td>
<td>0.066</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Only GR</td>
<td>0.122</td>
<td>0.703</td>
<td>0.250</td>
<td>0.584</td>
<td>0.427</td>
<td>0.410</td>
<td>0.591</td>
<td>0.241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Only GR</td>
<td>0.275</td>
<td>0.559</td>
<td>0.482</td>
<td>0.354</td>
<td>0.654</td>
<td>0.175</td>
<td>0.743</td>
<td>0.078</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Full</td>
<td>0.128</td>
<td>0.697</td>
<td>0.261</td>
<td>0.572</td>
<td>0.439</td>
<td>0.397</td>
<td>0.600</td>
<td>0.232</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Full</td>
<td>0.243</td>
<td>0.590</td>
<td>0.438</td>
<td>0.399</td>
<td>0.620</td>
<td>0.211</td>
<td>0.724</td>
<td>0.099</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex baseline</td>
<td>0.122</td>
<td>0.703</td>
<td>0.255</td>
<td>0.579</td>
<td>0.431</td>
<td>0.405</td>
<td>0.593</td>
<td>0.239</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex baseline</td>
<td>0.239</td>
<td>0.594</td>
<td>0.446</td>
<td>0.391</td>
<td>0.629</td>
<td>0.202</td>
<td>0.730</td>
<td>0.093</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Only GR</td>
<td>0.103</td>
<td>0.720</td>
<td>0.217</td>
<td>0.614</td>
<td>0.387</td>
<td>0.449</td>
<td>0.559</td>
<td>0.275</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Only GR</td>
<td>0.207</td>
<td>0.624</td>
<td>0.400</td>
<td>0.436</td>
<td>0.596</td>
<td>0.236</td>
<td>0.713</td>
<td>0.111</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Full</td>
<td>0.108</td>
<td>0.716</td>
<td>0.226</td>
<td>0.606</td>
<td>0.397</td>
<td>0.440</td>
<td>0.566</td>
<td>0.268</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Full</td>
<td>0.183</td>
<td>0.646</td>
<td>0.361</td>
<td>0.476</td>
<td>0.558</td>
<td>0.276</td>
<td>0.690</td>
<td>0.137</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex baseline</td>
<td>0.089</td>
<td>0.733</td>
<td>0.195</td>
<td>0.635</td>
<td>0.357</td>
<td>0.480</td>
<td>0.530</td>
<td>0.305</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex baseline</td>
<td>0.162</td>
<td>0.666</td>
<td>0.342</td>
<td>0.494</td>
<td>0.546</td>
<td>0.289</td>
<td>0.683</td>
<td>0.143</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Only GR</td>
<td>0.074</td>
<td>0.746</td>
<td>0.164</td>
<td>0.665</td>
<td>0.313</td>
<td>0.522</td>
<td>0.489</td>
<td>0.347</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Only GR</td>
<td>0.138</td>
<td>0.688</td>
<td>0.297</td>
<td>0.537</td>
<td>0.504</td>
<td>0.331</td>
<td>0.659</td>
<td>0.170</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Full</td>
<td>0.077</td>
<td>0.744</td>
<td>0.169</td>
<td>0.659</td>
<td>0.320</td>
<td>0.515</td>
<td>0.495</td>
<td>0.340</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Full</td>
<td>0.123</td>
<td>0.702</td>
<td>0.266</td>
<td>0.568</td>
<td>0.465</td>
<td>0.372</td>
<td>0.629</td>
<td>0.202</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex baseline</td>
<td>0.060</td>
<td>0.759</td>
<td>0.139</td>
<td>0.688</td>
<td>0.274</td>
<td>0.560</td>
<td>0.445</td>
<td>0.391</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex baseline</td>
<td>0.108</td>
<td>0.716</td>
<td>0.249</td>
<td>0.584</td>
<td>0.451</td>
<td>0.386</td>
<td>0.620</td>
<td>0.211</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Only GR</td>
<td>0.050</td>
<td>0.767</td>
<td>0.115</td>
<td>0.710</td>
<td>0.235</td>
<td>0.597</td>
<td>0.403</td>
<td>0.434</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Only GR</td>
<td>0.091</td>
<td>0.731</td>
<td>0.212</td>
<td>0.619</td>
<td>0.406</td>
<td>0.431</td>
<td>0.588</td>
<td>0.245</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex CF Full</td>
<td>0.053</td>
<td>0.765</td>
<td>0.121</td>
<td>0.704</td>
<td>0.244</td>
<td>0.589</td>
<td>0.412</td>
<td>0.425</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no sex CF Full</td>
<td>0.080</td>
<td>0.741</td>
<td>0.186</td>
<td>0.644</td>
<td>0.364</td>
<td>0.473</td>
<td>0.549</td>
<td>0.286</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aCF refers to counter-factual, GR refers to gender ratio. Each cell gives the probability of matching in sex or no sex markets based on an individual's grade and possible partner grade. $P_w^m$ is the probability of matching for a man looking for a woman, $P_m^w$ is the probability of matching for a woman looking for a man.
Table 12: Racial Gap in the Probability of Sex Conditional Matching Under Various Market Conditions

| Aggregate School | Same Grade and Race $P$(Sex|Match)$^a$ | 9    | 10    | 11    | 12    |
|------------------|----------------------------------------|------|-------|-------|-------|
| White            | 0.313 0.388 0.454 0.507               |      |       |       |       |
| Black            | 0.381 0.492 0.578 0.634               |      |       |       |       |
| Difference       | -0.067 -0.104 -0.125 -0.127           |      |       |       |       |

Changing:

Only Gender Ratios

| White            | 0.319 0.399 0.466 0.518               |      |       |       |       |
| Black            | 0.381 0.493 0.579 0.634               |      |       |       |       |
| Difference       | -0.062 -0.094 -0.113 -0.116           |      |       |       |       |

Gender Ratios and Male Past Sex

| White            | 0.313 0.387 0.452 0.506               |      |       |       |       |
| Black            | 0.349 0.431 0.496 0.536               |      |       |       |       |
| Difference       | -0.036 -0.044 -0.043 -0.031           |      |       |       |       |

Gender Ratios and All Past Sex

| White            | 0.313 0.387 0.453 0.506               |      |       |       |       |
| Black            | 0.322 0.398 0.462 0.513               |      |       |       |       |
| Difference       | -0.010 -0.011 -0.010 -0.007           |      |       |       |       |

$^a$Gap is measured as $P$(sex|match,white with white) - $P$(sex|match,black with black). Counterfactual policy simulation changes the black gender ratios to match those of whites in three stages, changing the grade-specific gender ratio, the past-sex distribution for black males, and the past-sex distribution for both black females and black males.
Table 13: Varying First-Stage Assumptions$^a$

<table>
<thead>
<tr>
<th>Mean $\phi_{sex}$</th>
<th>% Never Matched Removed</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>25</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$P^{out}_{match}$</td>
<td>$P^{out}_{match}$</td>
<td>$P^{out}_{match}$</td>
<td></td>
</tr>
<tr>
<td>Male, No Equilibrium</td>
<td>1</td>
<td>0.66</td>
<td>1</td>
<td>0.66</td>
</tr>
<tr>
<td>Female, No Equilibrium</td>
<td>0.581</td>
<td>0.768</td>
<td>0.784</td>
<td>0.746</td>
</tr>
<tr>
<td></td>
<td>0.332</td>
<td>0.220</td>
<td>0.225</td>
<td>0.246</td>
</tr>
</tbody>
</table>

Mean Stated Preference

<table>
<thead>
<tr>
<th>Mean Stated Preference</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.574</td>
<td>0.564</td>
<td>0.594</td>
</tr>
<tr>
<td>Female</td>
<td>0.341</td>
<td>0.332</td>
<td>0.355</td>
</tr>
</tbody>
</table>

$^a$Number of observations is different for each model. Removing never matched involves shrinking the estimation sample by random removal on never-matched individuals, and shrinking the aggregate number of searching men and women with the probability of never-matching estimated with a logit at the type-school level, separately for men and women. Decreasing the probability of matching outside the school also shrinks the estimation sample the and aggregate number of searching men and women in a similar fashion.