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

Peter Arcidiacono
Duke University
& NBER

Andrew Beauchamp
Boston College

Marjorie McElroy
Duke University
& NBER

June 2013

Abstract: We develop a two-sided directed search model of relationship formation which can be used to disentangle male and female preferences over partner characteristics and different relationship terms from only a cross-section of observed matches. Individuals 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. 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 the structural model match subjective data on whether sex would occur in one's ideal relationship. The equilibrium result shows that some women would ideally not have sex, but do so out of matching concerns; the reverse is true for men. 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, Chicago, 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

Next to drinking alcohol, sex is the most common risky behavior of teens.[1] While numerous studies try to measure the impact of policy or technological changes on sexual behaviors of teens[2] and others try to measure the effect of sexual activities on educational and health outcomes,[3] to our knowledge there are no economic studies of a fundamental tradeoff faced by teens: the tradeoff between being in a romantic relationship at all and the inclusion of sex in a relationship. We model and estimate this tradeoff within the confines of the best available data on teen relationships and sexual behaviors. We show that in equilibrium some women have sex out of matching concerns. Further, as men become more scare, this effect is accentuated.

To uncover differences in gender-specific preferences, we specify a model of two-sided directed search. Utilities in a relationship depend on partner characteristics, the terms of the relationship, and on past sexual experience.[4] Individuals of a discrete type on one side of the market target their search toward potential partners with particular characteristics and also target their search based on the terms of the relationship (sex or no sex in our case).[5] Given the supplies of individuals of each type on both sides of the market, the ex ante yield from targeting particular partner characteristics and terms depends on the associated probability of matching. The probability of matching in turn is endogenously determined by the choices of individuals on one's own side of the market (rivals) and by all individuals on the other side (prospects). Changes in the supply of men and women of different characteristics

[1]The Youth Risk Behavior Surveillance System (YRBS) of the Centers for Disease Control tracks the major risk taking behaviors of a national representative sample of ninth through twelfth graders. Summary statistics for 2011 show that more teens have had sex (47%) than ever used marijuana (40%), other illicit drugs, tobacco (44%), or attempted suicide. Only the behavior of ever having a drink of alcohol (71%) outranked sex. http://www.cdc.gov/healthyyouth/yrbs/pdf/us_overview_yrbs.pdf

[2]Akerlof, Yellen and Katz (1996) discuss how abortion changed sexual particiaption; Girma and Paton (2011) study emergency contraception availability and unprotected sex.

[3]See Sabia and Rees (2008) and Sabia and Rees (2009).

[4]Arcidiacono, Khwaja and Ouyang (2012) provide evidence that once adolescent females become sexually active, they generally remain so.

[5]This formulation of characteristics and terms builds on work by Dagsvik (2000) who 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. Willis (1999) also presents an equilibrium model where the terms are joint and single parenthood, showing that sex ratio changes can generate equilibria where women raise children as single mothers. Recent work by Salani and Galichon (2012) addresses matching based on unobserved characteristics under transferable utility.

shift the equilibrium distribution of terms and partner characteristics.

Thus our model lies in a middle ground between models in which the assignment to partners is efficient[6] and search models in which the decision to accept a match is endogenous but the probability of meeting a particular type of partner is exogenous.[7] Our model also is somewhat analogous to wage posting models in which workers know the terms of the relationship (the wage) they would receive from each potential employer.[8]

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.[9,10] Men and women target their searches not only based upon the multidimensional characteristics of the partner but also on the terms of the relationship.[11] 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.[12] 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 elas-

[6]See Choo and Siow (2006), Siow (2009) and Fox (2010).

[7]See, for example, Brien, Lillard and Stern (2006), Ge (2011), Laufer and Gemici (2011), Rasul (2006), Richards-Shubik (2012), Wong (2003b), Wong (2003a) and Jacquemet and Robin (2013).

[8]See for example Burdett and Mortensen (1998)

[9]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.

[10]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.

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

[12]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.

ticity of substitution matching functions, we show that, as the gender ratio becomes more unfavorable, the individual becomes more likely to sacrifice relationship terms for a higher match probability.

The advantages of our modeling approach are three-fold. First, by linking choices over partners with outcomes, in equilibrium we are able to weight the different 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 grade 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.[13] In the transferable case, if the occurrence of a particular outcome is affected by the gender ratio, it is unclear how utilities 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 Add Health data. 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 the structural model show that men value sexual relationships relatively more than women. By simulating choices in the absence of matching concerns, we find that 31.6% of women and 61.8% of men would prefer to be in a sexual, as opposed to a nonsexual, relationship. These counterfactual choices

[13]Both Hitsch, Hortacsu and Ariely (2010) and Fisman, Iyengar, Kamenica and Simonson (2006) use observed choices of both men and women to back out preferences for partner characteristics. Fisman, Iyengar, Kamenica and Simonson (2006) observe choice in a speed-dating experiment and Hitsch, Hortacsu and Ariely (2010) observe them in an online-dating context. Chiappori, Oreffice and Quintana-Domeque (2009) recovers separate preferences for males and females for partner characteristics but not for relationship terms.

bear a striking resemblance to subjective reports by students found in Add Health. There, 34.6% of women and 58.3% 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. Conditional on matching, the data reveal that black females (and males) are substantially more likely to have sex than their white counterparts. We simulate changing the market black females face in order to understand the sources of the 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 small. 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 shows 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 the robustness of our results to different assumptions regarding searching outside the school as well as choosing not to search. Section 6 gives the counterfactuals. Section 7 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). Individuals know in advance their payoffs from different partner types and relationship terms. And, they may target less-preferred combinations of types and terms in order to have higher probabilities 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.[14] 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 im indicate the i-th member of type m.

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

[14] 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_m^{wr},

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

3. and an individual-specific preference term ϵ_{im}^{wr}.[15]

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

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

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

Taking logs yields:

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

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

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

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

We treat the ϵ_{im}^{wr}'s as observed only to the individual: only the distribution is known to the other participants in the market. We assume that the ϵ_{im}^{wr}'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, σ, 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, ϕ_m^{wr} then follows a multinomial logit form:

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

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 ϕ_m^{wr} and ϕ_w^{mr} give the probability of m-type men and w-type women who search in market $\{m, w, r\}$ which are also the market shares of searching men and women. 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:[16]

$$\begin{aligned} X_{mwr} &= A^*\left[\frac{(\phi_m^{wr}N_m)^\rho}{2}+\frac{(\phi_w^{mr}N_w)^\rho}{2}\right]^{\frac{1}{\rho}} \\ &= A\left[(\phi_m^{wr}N_m)^\rho+(\phi_w^{mr}N_w)^\rho\right]^{\frac{1}{\rho}} \end{aligned} \tag{5}$$

[16]For ease of exposition we are assuming an interior solution such that the number of matches produced is less than both the number of men and the number of women in the $\{w, m, r\}$ market. In practice, we nest the CES matching function into a Leontief function to constrain the number of matches to be less than the number of searching men and women:

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

where ρ 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_m^{wr} is given by:

$$\begin{aligned} P_m^{wr} &= \frac{X_{mwr}}{\phi_m^{wr} N_m} \\ &= \frac{A\left[(\phi_m^{wr} N_m)^{\rho} + (\phi_w^{mr} N_w)^{\rho}\right]^{\frac{1}{\rho}}}{\phi_m^{wr} N_m} \\ &= A\left[1 + \left(\frac{\phi_w^{mr} N_w}{\phi_m^{wr} N_m}\right)^{\rho}\right]^{\frac{1}{\rho}}. \end{aligned} \tag{6}$$

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 ϕ's, give the share of a particular set individuals who will search in a particular market. These ϕ's also affect the probabilities of matching, the P's. We rewrite equation (7) to make the dependence of P_m^{wr} on ϕ_m^{wr} and ϕ_w^{mr} explicit:

$$\phi_m^{wr} = \frac{\exp\left(\frac{\mu_m^{wr} + \ln[P_m^{wr}(\phi_m^{wr}, \phi_w^{mr})]}{\sigma}\right)}{\sum_{w'}\sum_{r'} \exp\left(\frac{\mu_w^{w'r'} + \ln\left[P_m^{w'r'}\left(\phi_m^{w'r'}, \phi_{w'}^{mr'}\right)\right]}{\sigma}\right)} \tag{7}$$

Since the market shares must sum to one for both men and women, equilibrium in our model is characterized by stacking the $(M-1) \times (W-1) \times (R-1)$ shares and solving for the fixed point. Since ϕ is a continuous mapping on a compact space, Brouwer's fixed point theorem guarantees that an equilibrium exists.[17] 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

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

cannot be a Pareto move.[18]

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 ϕ'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_w^m 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 $\mu_w^{mr'} - \mu_w^{mr} > \mu_m^{wr'} - \mu_m^{wr}$ then the following hold:*

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

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

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

In Proposition 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,

[18] Note that because the decisions are based on expected utility and probabilistic search the stability of the resulting matching will not hold.

for elasticities of substitution less than one ($\rho < 0$) and conditional on ϕ, the lack of substitutability between men and women in the market implies that, as G_w^m increases, larger changes in match probabilities for men (women) will occur where men are relatively more (less) abundant. The elasticity of the probability of matching with respect to G_w^m for men and women in the $\{m, w, r\}$ market conditional on ϕ are given by:

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

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

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

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

The number of matches is 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.[19]

3 Data and Descriptive Characteristics

We use data from Wave I of the National Longitudinal Survey of Adolescent Health.[20] The data include an in-school survey of almost 90,000 seventh to twelfth grade stu-

[19] On the other extreme is the Cobb Douglas case where $\rho \to 0$, implying the elasticity of substitution, σ, is one. Hence, if gender ratios do affect relationship terms then this is evidence that the matching function is not Cobb Douglas.

[20] 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.

dents at a randomly sampled set of 80 communities across the United States.[21] Attempts were made to have as many students as possible from each school fill out the survey during a school day. Questions consist mainly of individual data like age, race, and grade, with limited information on academics, extra-curricular activities and risky behavior. We use this sample to construct school level aggregates by observable characteristics, grade and race, 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 relationship 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. The definition that a "relationship" referred to from here on, consists of both 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.[22] The relationship history allows us to determine whether respondents 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 only 24% of matches.[23] 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. 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 5.3.[24]

[21] 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.

[22] 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.

[23] 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.

[24] From 14840 students between grades 9-12, we drop: schools discussed above (2622), schools with fewer than 10 reported students (146), unweighted data (996), missing data (106) and those matching outside the school (3771), giving 7141 valid observations for in-school searchers. Of those currently matched outside, roughly 48% were matched with a partner whose age or grade made them highly unlikely to be able to attend the same school at the time of survey. A presumably large and unknown fraction of these matches began as same school matches in the past. The school is still the primary matching market, despite our inability to observe same-school status in the past.

The Add Health data is nationally representative at the school-level and so is drawn from all types of schools. We focus on respondents who are in the 9th through 12th grades.[25] 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.[26]

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 688 matches where sex occurred and 561 matches where sex did not occur, while women reported 551 matches with sex and 532 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.[27] This also ensures that our estimated results on differences in male and female preferences for sex are driven not by differences in self-reporting.

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.[28] 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. Current sexual participation is also higher for black males and females than the other race/ethnicity groups.

Unweighted data are dropped because they come from schools where we do not observe the aggregate data.

[25]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.

[26]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 most relevant junior high or middle school for a high school, the relevant 9th grade observations for these six schools were not sampled, but rather a small "feeder" school.

[27]Male reports were much noisier than female reports with respect to their correlation with the gender-ratio, suggesting either female willingness to truthfully report sex varies systematically and robustly with the gender ratio, or what we argue is more likely, men over-report sex.

[28]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.

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.[29] 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 (over 95% want a relationship as defined), preferences for sex are not. While 60% of men would prefer to have sex, the fraction of women who prefer to have sex is only 36%. Preferences for sex rise with age. Even with this rise, comparing the sex preferences for women of a particular grade with the sex preferences for men of another grade shows stronger male preferences for sex. 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.

To investigate this further, Table 3 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 then men who don't want to have sex. The second row shows 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 within 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 12 points more likely to have sex conditional on wanting to have

[29] 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.

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 4 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.[30] 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 5 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/ethnicity 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.[31] 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

[30]The data presented here are from the searching sample, which is adjusted via a first-stage estimate of the probability of searching outside the school. This accounts for why the fraction-of-the-sample declines with grade.

[31]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.

Wong (2003b), who argues a marriage taboo dramatically influences the frequency of cross-race matching among black males. Black men 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 , 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 ratio of female to male students for each grade-race pairing.[32] Table 6 shows that there is a substantial amount of variation in the gender ratio, particularly among blacks.[33] Breaking out the gender ratio along different dimensions (race, and grade-race groupings) spreads the initially condensed distribution.[34]

The bottom panel of Table 6 shows the probability of having sex conditional matching conditional on the fraction female being above the 75th percentile or below the 25th percentile. The first two rows show the cases when gender ratio is measured using the whole school and then using only those of the same race. In both cases, a ratio of females to males 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 ratio of females to males among 12th grade whites. 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 cell is higher the probability of sex in the relationship is more likely: when women face more competition for partners, more sex results.

To further investigate the role of competition in determining relationship terms,

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

[33] This dispersion is even more pronounced for Hispanic and other-race students.

[34] 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.

Table 7 presents marginal effects from a probit estimated on the probability of sex conditional on matching for women. We again use the ratio of females to males in the partner's grade-race cell 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 school fixed effects. 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 associated with the choice of partner characteristics.

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 having sexual intercourse in their current match, 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, (3) unreported partner characteristics, and (4) aggregating to the school-level from our in-home sample. We discuss how we deal with each of these issues in Appendix B.

4.1 Utility

Rather than having separate μ's (utilities) for every type of relationship, we put some structure on the utility function. Denote the grade associated with an m-type man as $G_m \in \{1, 2, 3, 4\}$. When a man searches for an w-type woman, the grade of the partner is PG_w. Similarly, $R_m \in \{1, 2, 3, 4\}$ gives the race of an m-type man with the corresponding race of the potential w-type partner given by PG_w. We specify the utility of a non-sexual relationship as a function of the partner's grade and race as well as whether the partner is in the same grade as the searching individual, $SG_{mw} = I(G_m = PG_w)$ where I is the indicator function, and the same race, $SR_{mw} = I(R_m = PR_m)$.

Denoting searching in the no-sex market by $r = 1$, we formulate the deterministic part of utility for men and women matching in the no-sex market as:

$$\mu_{mw1} = \alpha_1 SG_{mw} + \alpha_2 PG_w + \alpha_3 SR_{mw} + \sum_{j=1}^{4} I(PR_w = j)\alpha_{4j} \quad (8)$$

$$\mu_{wm1} = \alpha_1 SG_{mw} + \alpha_5 PG_m + \alpha_3 SR_{mw} + \sum_{j=1}^{4} I(PR_m = j)\alpha_{6j} \quad (9)$$

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 in 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. We also include a grade profile which captures the transition process into sexual activity (e.g. social, biological and other factors changing as adolescents age). 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:

$$\mu_{mw2}(PS_{im}) = \mu_{mw1} + \alpha_7 + \alpha_8 PS_{im} + \alpha_{12} G \quad (10)$$
$$\mu_{wm2}(PS_{iw}) = \mu_{wm1} + \alpha_{11} + \alpha_8 PS_{iw} + \alpha_{13} G \quad (11)$$

Although men an women may differ in their preferences for sex, the effect of past sex is constrained to be the same for men and women.[35]

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 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. Since women report their partner characteristics we only require a sample of women (matched and unmatched) along with the aggregate type-distributions of men and women to estimate our model.[36]

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

[35] 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.

[36] Women reported partner grade and race; partner's past sexual experiences were not reported, so we estimate this probability conditional on grade and race from the male reports.

the ith woman of type w is then given by:

$$\begin{aligned} L_{iw}(\theta) &= I(y_{iw}=1)\left[\sum_m \sum_r I(d_{iw}=\{m,r\})\left(\ln\left[\phi_w^{mr}(\theta,\mathbf{N},PS_{iw})\right]+\ln\left[P_w^{mr}(\theta,\mathbf{N})\right]\right)\right] \\ &\quad + I(y_{iw}=0)\ln\left[\sum_m \sum_r \phi_w^{mr}(\theta,\mathbf{N},PS_{iw})\times(1-P_w^{mr}(\theta,\mathbf{N}))\right] \end{aligned} \tag{12}$$

Note that the probability of matching is not affected by past sex except through the search probabilities. The ϕ terms are the equilibrium probabilities of searching in each market for each type individual.

The likelihood described so far was for a generic school. Denote the schools in the data by $s \in \{1,\dots,S\}$. 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:

$$\hat{\theta} = \arg\max_{\theta}\left(\sum_s \sum_w \sum_{i=1}^{N_w^s} L_{iw}^s(\theta)\right)$$

where a fixed point in the search probabilities for men and women is solved at each iteration.

5 Results

The estimates of the structural model are presented in Table 8.[37] Key to disentangling male and female preferences given observed matches is the effect of the different gender ratios on the search decisions. These gender ratios manifest themselves through their effect on the probability of matching. The parameters of the matching function, ρ and A, are identified through variation in matches across schools with different gender ratios and the overall match rate respectively.[38] The estimates of ρ are significant and

[37] We also estimated models with a nested logit structure on race and sex separately, and a product-differentiation logit (e.g. a nested logit with overlapping nests), nesting parameters were not significantly different from one, and we could not reject the null that the logit model fit the data best. Because we estimate the variance of the logit errors, the interpretation of the correlation parameters is non-standard.

[38] In a 2-market model with only male and female preferences for sex, ρ 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

negative, ruling out the Cobb-Douglas matching model and confirming that gender ratios do affect the likelihood of observing particular matches.

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 preferences for sex than females. Those who have had sex in the past also have a much stronger preference for sex in the present, and as adolescents age their utility from sex increases. Despite the negative preferences for sex, it is not the case there is no incentive for either men or women to engage in sex. On the contrary the value of matching with a given partner must pass a higher valuation threshold in order for sex to be preferred if an individual is sexually inexperienced. This threshold for sex then drops if one has had sex in the past. Thus the model predicts individuals should initiate sex with higher value partners, than those with whom they subsequently would be willing to have sex with. The grade profiles show that the female value of having sex with older partners is more than twice that of males.

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 men have a slight preference for younger partners. 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, and individuals have values associated with matching within their own race.

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

counting the gender ratio overall). Thus the search friction A is identified by including multiple schools. Cite See Hsieh (2012) for recent work on identification in models with different male and female preferences within a single market.

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 good job of matching the elicited preferences for sex, which we report in two ways, conditional on wanting a match and unconditionally (as reported in Table 9). The elicited preferences show 34-36% of women prefer sex to 31.6% of women predicted by the model, while for men the elicited preference for sex is between 58% and 60% compared to a model prediction of 61.8%. The model predictions and self-reports both show a higher preference for sex among blacks relative two whites. The predicted male-race profile is extremely close the subjective reports, while the model under-predicts female preferences relative to the subjectives. The grade profiles are also similar. Beginning around 50%, male preferences for sex increase with grade reaching above 70% by 12th grade. For females we see growth from around 20% to roughly 50% between 9th and 12th grade in both stated and predicted preferences.

5.2 Equilibrium Match Probabilities

Table 10 presents the estimated equilibrium match probabilities for whites.[39] 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_w^m) . 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_w^m in the first row of cells give the comparable eight probabilities for ninth grade women. The grade-matching patterns are driven by the equilibrium combination of utility parameters for partner and own grade for men and women and probability of matching. As expected, ninth grade males – in both the sex and no sex markets – see the lowest match probabilities of any group. At the extreme, all females can always match with a ninth-grade male in the sex market, but preferences for partner grade lead them to search elsewhere. Females see a higher probability of

[39]The equilibrium includes individuals from all 32 groups; we present only the subset of probabilities for whites in Table 10

matching in the no-sex market when they are older, but that gain does not translate into a higher probability of matching in the sex market, again because of preferences. This is due to female preferences for sex at older ages along with male preferences for younger partners, which work against the older female probability of matching in the sex market. For males, the probability of matching declines with female age in the no-sex market because of a female preference for partner grade (4.750 in Table 8) and the male dis-utility for partner grade (-.668 in Table 8). 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 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:

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

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

$$\Pr(w, r|m, d \neq 0) = \phi_m^{wr} = \frac{\exp\left(\frac{\mu_m^{wr} + \ln(P_m^{wr})}{\sigma}\right)}{\sum_{w'}\sum_{r'}\exp\left(\frac{\mu_m^{w'r'} + \ln\left(P_m^{w'r'}\right)}{\sigma}\right)}, \tag{14}$$

which is the probability in equation (7) that we used to form our likelihoods, assuming everyone who did not match in the outside market searched within the school.

When the probability of matching in the outside market is not one, the IIA property still holds, but we no longer have a good measure of the number of searchers in the outside market. We allow for the probability of matching in the outside market to be less than one by assuming that match rates are the same for all those who search outside the school. Namely, we know the characteristics of those who matched in the outside market. If 10% of freshmen girls and 16% of senior girls match in the outside market, for a candidate match rate of 80%, the implied freshmen girl search rate would 12.5% in the outside market with the corresponding search rates for senior girls being 20%. We then scale down the number of non-matched individuals with particular characteristics given the implied outside-market search rates to determine the number of women searching within the school. For unmatched individual's in the sample we randomly assign status as outside-searchers, based on their type-school-specific probability of searching outside, which combines data on matching outside with an assumption on the match rate. We then drop these outside-searchers.

More formally, denote N_m^0 as the number of m-type men who matched in the outside market. N_m was defined as the number of searching men of type m in the school and was formed as the number of m-type men minus N_m^0. With the probability of matching in the outside market given by P_m^0, instead of forming N_m by subtracting off N_m^0 for the population, we subtract off N_m^0/P_m^0. Note that the utilities of searching in the outside market are allowed to vary conditional on school and characteristics. The only restriction is that, conditional on searching in the outside market, the match rate is the same. We need this in order to characterize the number of those searching inside the school.[40]

Results for the estimated model under different match rate assumptions are presented in Table 11, with the first column repeating the results of the original model. The top rows of Table 11 show the mean choice probability for choosing sex without equilibrium influences. For reasonable match rates, we continue to see the same patterns of women preferring sex relative to men. However, as the match rate falls, ρ moves closer to zero, which is the Cobb-Douglas case.[41]

[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_m^0, F_w^0, 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_m^0, but estimate it as $N_m^0 = \hat{F}_m^0 N_m$.

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

As we move across the table, the composition of the sample changes. This is because we drop individual's who we randomly assign as outside-searchers, in proportion to individuals in their school-type who matched outside. 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 to 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 results are presented in the second and third sets of columns Table 11.

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. With a reasonable friction in the outside market (66% probability of matching) the predicted preferences do not fit the subjectives any better than the baseline model. 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.

6 Counterfactual Environments for Black Women

Our structural model permits us to perform counterfactual simulations. We aggregate our sample of 73 Add Health schools to create one large, representative school.[42] Black women in relationships are roughly 10 percentage points more likely to have sex than white women. 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

[42]The constant returns-to-scale matching technology makes the number of individuals irrelevant.

out: (i) as compared to whites, among 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 also 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 (iii) change the distribution of past sex among black women to match that of white women.[43] Examining 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 12 we present the aggregate school simulated matching probabilities for black men (P_m^w) and black women (P_w^m) (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, whiles 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 (1.00) than her male classmate (0.11); and her male classmate can triple his probability of matching (to 0.33) 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),[44] but not uniformly. Consistent with

[43]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.

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

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. 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 13 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 same-race matches only.[45] The "racial gap" compares probabilities of a match in the sex market *conditional on matching*. Table 13 reports the difference between these conditional probabilities for whites matched with whites minus that for blacks matched with blacks. The baseline shows the 9 to 10 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 (roughly 15% for 9th graders). Additionally adjusting the past-sex distribution for black men (ii) yields a much larger effect, between 65% and 75% for 10th-12thers graders. 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 preferences over partner race and partner grade. The full simulations show that a remarkable 75%-100% of the racial gap in sexual participation between

[45] Results look similar within a partner grade category.

similar black and white women can be attributed to the differences in the matching markets they face.

7 Conclusion

The contribution of this paper is two-fold. First, we show how a directed search model can disentangle male and female preferences for different 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.

More generally, because changes in the supplies of various types of men and women alters the equilibrium match distribution, such changes have the potential to deeply effect the wellbeing of individuals – changing who enters unions of many types, who does not enter at all, the characteristics of partners for those who do match, the intra-union distribution of the surplus from unions and ultimately the health and wellbeing of most everyone in current and future generations. These important links are reflected in a growing literature on the effects of imbalances and changes in sex ratios within both developed and developing countries. Many of these studies focus on the effects of changing sex ratios on who does and does not enter unions and with whom those who enter match.[46] Other studies address the partial equilibrium effects on the intra-union distribution of the match surplus.[47] But the methods in these studies cannot reveal the separate preferences of men and women and how the tradeoffs they face relate to the supplies of various types of men and women. Our model and empirics take a step in that direction.

References

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

Akerlof, G.A., J.L. Yellen, and M.L. Katz, "An Analysis of Out-of-Wedlock Childbearing in the United States," *Quarterly Journal of Economics*, May 1996, *11* (2), 277–317.

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

Arcidiacono, Peter, Ahmed Khwaja, and Lijing Ouyang, "Habit Persistence and Teen Sex: Could Increased Access to Contraception Have Unintended Consequences for Teen Pregnancies?," *Journal of Business Economic Statistics*, 2012, *30* (2), 312–325.

Brien, Michael J., Lee A. Lillard, and Steven Stern, "Cohabitation, Marriage, And Divorce In A Model Of Match Quality," *International Economic Review*, 05 2006, *47* (2), 451–494.

Bruze, Gustaf, Michael Svarer, and Yoram Weiss, "The Dynamics of Marriage and Divorce," IZA Discussion Papers 6379, Institute for the Study of Labor (IZA) February 2012.

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

[46]Examples include Wilson (1987), Willis (1999), Abramitzky, Delavande and Vasconcelos (2010), Rose (2004)

[47]See Chiappori, Oreffice and Quintana-Domeque (2009) and Bruze, Svarer and Weiss (2012).

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

_ **, Sonia Oreffice, and Climent Quintana-Domeque**, "Fatter Attraction: Anthropometric and Socioeconomic Characteristics in the Marriage Market," IZA Discussion Papers 4594, Institute for the Study of Labor (IZA) November 2009.

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

Dagsvik, John K, "Aggregation in Matching Markets," *International Economic Review*, February 2000, *41* (1), 27–57.

Del Boca, Daniela and Christopher Flinn, "Endogenous household interaction," *Journal of Econometrics*, 2012, *166* (1), 49–65.

_ **and** _ , "Household Behavior and the Marriage Market," Technical Report 2012.

Dupuy, Arnaud and Alfred Galichon, "Personality Traits and the Marriage Market," IZA Discussion Papers 6943, Institute for the Study of Labor (IZA) October 2012.

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

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

Ge, Suqin, "Womens College Decisions: How Much Does Marriage Matter?," *Journal of Labor Economics*, 2011, *29* (4), pp. 773–818.

Girma, S. and David Paton, "The Impact of Emergency Birth Control on Teen Pregnancy and STIs," *Journal of Health Economics*, 2011, *Forthcoming.*

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

Hsieh, Yuwei, "Understanding Mate Preferences from Two-Sided Matching Markets: Identification, Estimation and Policy Analysis," Working Paper, University of Southern California 2012.

Jacquemet, Nicolas and Jean-Marc Robin, "Assortative matching and search with labour supply and home production," CeMMAP working papers CWP07/13, Centre for Microdata Methods and Practice, Institute for Fiscal Studies March 2013.

Laufer, Steve and Ahu Gemici, "Marriage and Cohabitation," 2011 Meeting Papers 1152, Society for Economic Dynamics 2011.

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

Lundberg, Shelly and Robert A Pollak, "Separate Spheres Bargaining and the Marriage Market," *Journal of Political Economy*, December 1993, *101* (6), 988–1010.

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

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

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

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

Rasul, Imran, "Marriage Markets and Divorce Laws," *Journal of Law, Economics and Organization*, April 2006, *22* (1), 30–69.

Richards-Shubik, Seth, "Peer Effects in Sexual Initiation: Separating Demand and Supply Mechanisms," Working Paper, Carnegie Mellon University 2012.

Rose, Elaina, “Education and Hypergamy in Marriage Markets,” Working Paper 353330, University of Washington 2004.

Sabia, Joseph J. and Daniel I. Rees, “The effect of adolescent virginity status on psychological well-being,” *Journal of Health Economics*, 2008, *27* (5), 1368 – 1381.

_ **and** _ , “The Effect of Sexual Abstinence on Females’ Educational Attainment,” *Demography*, 2009, *46* (4), 695715.

Salani, Bernard and Alfred Galichon, “Cupid’s Invisible Hand: Social Surplus and Identification in Matching Models,” Technical Report 2012.

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

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

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

Wong, Linda Y., “Structural Estimation of Marriage Models,” *Journal of Labor Economics*, July 2003, *21* (3), 699–728.

_ , “Why so only 5.5% of Black Men Marry White Women?,” *International Economic Review*, 2003, *44* (3), 803–826.

Appendix A

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

$$\begin{aligned} &\mu_w^{mr'} + \log(P_w^{mr'}) - \mu_w^{mr} - \log(P_w^{mr}) + \log(P_m^{wr'}) - \log(P_m^{wr}) > \\ &\quad \mu_m^{wr'} + \log(P_m^{wr'}) - \mu_m^{wr} - \log(P_m^{wr}) + \log(P_w^{mr'}) - \log(P_w^{mr}). \end{aligned} \tag{15}$$

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

$$\frac{e^{\mu_w^{mr'} + \log(P_w^{mr'})}}{e^{\mu_w^{mr} + \log(P_w^{mr})}} e^{\log(P_m^{wr'}) - \log(P_m^{wr})} > \frac{e^{\mu_m^{wr'} + \log(P_m^{wr'})}}{e^{\mu_m^{wr} + \log(P_m^{wr})}} e^{\log(P_w^{mr'}) - \log(P_w^{mr})} \tag{16}$$

or:

$$\frac{\phi_w^{mr'}}{\phi_w^{mr}}\frac{P_m^{wr'}}{P_w^{mr'}} > \frac{\phi_m^{wr'}}{\phi_m^{wr}}\frac{P_m^{wr}}{P_w^{mr}}. \tag{17}$$

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

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

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

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

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

$$\frac{\phi_w^{mr'}}{\phi_w^{mr}}\left[\left(\frac{\phi_w^{mr'}N_w}{\phi_m^{wr'}N_m}\right)^{\rho}\right]^{1/\rho} > \frac{\phi_m^{wr'}}{\phi_m^{wr}}\left[\left(\frac{\phi_w^{mr}N_w}{\phi_m^{wr}N_m}\right)^{\rho}\right]^{1/\rho}, \tag{17}$$

which further simplifies to:

$$\left(\frac{\phi_w^{mr'}}{\phi_w^{mr}}\right)^2 > \left(\frac{\phi_m^{wr'}}{\phi_m^{wr}}\right)^2, \tag{17}$$

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:

$$\frac{\phi_w^{mr}}{\phi_m^{wr}} < \frac{\phi_w^{mr'}}{\phi_m^{wr'}} \tag{17}$$

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

$$\left(\frac{\phi_w^{mr}N_w}{\phi_m^{wr}N_m}\right)^{1/\rho} > \left(\frac{\phi_w^{mr'}N_w}{\phi_m^{wr'}N_m}\right)^{1/\rho} \tag{17}$$

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

$$\left[1+\left(\frac{\phi_w^{mr}N_w}{\phi_m^{wr}N_m}\right)^{1/\rho}\right]^{\rho} < \left[1+\left(\frac{\phi_w^{mr'}N_w}{\phi_m^{wr'}N_m}\right)^{1/\rho}\right]^{\rho} \tag{17}$$

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

To evaluate claim (c), we use the implicit function theorem coupled with Cramer's rule. We show the case when there is only one type of man and one type of woman with two relationship types, r and r'. For ease of notation, we then denote $G = G_r^m$. Our proof, however, holds in the general case due to the independence of irrelevant alternatives associated with the Type I extreme value errors. Note that the definitions of search probabilities imply that, in equilibrium the log odds ratios for women satisfy:

$$\ln(\phi_w^{mr'})-\ln(1-\phi_w^{mr'}) \equiv \ln(\mu_w^{mr'})-\ln(\mu_w^{mr})+\ln\left[1+\left(\frac{\phi_m^{wr'}G}{\phi_w^{mr'}}\right)^{\rho}\right]-\ln\left[1+\left(\frac{(1-\phi_m^{wr'})G}{1-\phi_w^{mr'}}\right)^{\rho}\right] \tag{17}$$

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

$$\begin{aligned} F_1(\phi_w^{mr'}, \phi_m^{wr'}, G) \equiv{}& \ln(\phi_w^{mr'}) - \ln(1-\phi_w^{mr'}) - \ln(\mu_w^{mr'}) + \ln(\mu_w^{mr}) \\ &- \ln\left[1+\left(\frac{\phi_m^{wr'}G}{\phi_w^{mr'}}\right)^{\rho}\right] + \ln\left[1+\left(\frac{(1-\phi_m^{wr'})G}{1-\phi_w^{mr'}}\right)^{\rho}\right] \qquad (18) \\ F_2(\phi_w^{mr'}, \phi_m^{wr'}, G) \equiv{}& \ln(\phi_m^{wr'}) - \ln(1-\phi_m^{wr'}) - \ln(\mu_m^{wr'}) + \ln(\mu_m^{wr}) \\ &- \ln\left[1+\left(\frac{\phi_w^{mr'}}{\phi_m^{wr'}G}\right)^{\rho}\right] + \ln\left[1+\left(\frac{1-\phi_w^{mr'}}{(1-\phi_m^{wr'})G}\right)^{\rho}\right] \qquad (18) \end{aligned}$$

which can equivalently be expressed as:

$$\begin{aligned} F_1(\phi_w^{mr'}, \phi_m^{wr'}, G) \equiv{}& 2\ln(\phi_w^{mr'}) - 2\ln(1-\phi_w^{mr'}) - \ln(\mu_w^{mr'}) + \ln(\mu_w^{mr}) \qquad (18) \\ &- \ln\left[\left(\phi_w^{mr'}\right)^{\rho} + \left(\phi_m^{wr'}G\right)^{\rho}\right] + \ln\left[(1-\phi_w^{mr})^{\rho} + \left((1-\phi_m^{wr'})G\right)^{\rho}\right] \\ F_2(\phi_w^{mr'}, \phi_m^{wr'}, G) \equiv{}& 2\ln(\phi_m^{wr'}) - 2\ln(1-\phi_m^{wr'}) - \ln(\mu_m^{wr'}) + \ln(\mu_m^{wr}) \qquad (18) \\ &- \ln\left[\left(\phi_m^{wr'}\right)^{\rho} + \left(\frac{\phi_w^{mr'}}{G}\right)^{\rho}\right] + \ln\left[(1-\phi_m^{wr'})^{\rho} + \left(\frac{1-\phi_w^{mr'}}{G}\right)^{\rho}\right] \end{aligned}$$

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

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

which can be expressed as:

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

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

$$\frac{\partial F_1}{\partial G} = G^{-1} \left(\left[\left(\frac{1 - \phi_w^{mr'}}{(1 - \phi_m^{wr'})G} \right)^{\rho} + 1 \right]^{-1} - \left[\left(\frac{\phi_w^{mr'}}{\phi_m^{wr'} G} \right)^{\rho} + 1 \right]^{-1} \right) \tag{18}$$

$$\frac{\partial F_2}{\partial G} = G^{-1} \left(\left[\left(\frac{\phi_m^{wr'} G}{\phi_w^{mr'}} \right)^{\rho} + 1 \right]^{-1} - \left[\left(\frac{(1 - \phi_m^{wr'})G}{1 - \phi_w^{mr'}} \right)^{\rho} + 1 \right]^{-1} \right) \tag{19}$$

$$\tag{20}$$

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

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

$$\frac{\partial F_1}{\partial \phi_w^{mr'}} = \frac{(\phi_w^{mr'})^{\rho} + 2(\phi_m^{wr'} G)^{\rho}}{\phi_w^{mr'} \left((\phi_w^{mr'})^{\rho} + (\phi_m^{wr'} G)^{\rho} \right)} + \frac{(1 - \phi_w^{mr'})^{\rho} + 2((1 - \phi_m^{wr'})G)^{\rho}}{(1 - \phi_w^{mr'}) \left((1 - \phi_w^{mr'})^{\rho} + ((1 - \phi_w^{mr'})G)^{\rho} \right)} > 0 \tag{21}$$

$$\frac{\partial F_1}{\partial \phi_m^{wr'}} = - \frac{(\phi_m^{wr'} G)^{\rho}}{\phi_m^{wr'} \left((\phi_w^{mr'})^{\rho} + (\phi_m^{wr'} G)^{\rho} \right)} - \frac{((1 - \phi_m^{wr'})G)^{\rho}}{(1 - \phi_m^{wr'}) \left((1 - \phi_w^{mr'})^{\rho} + ((1 - \phi_w^{mr'})G)^{\rho} \right)} < 0 \tag{22}$$

$$\frac{\partial F_2}{\partial \phi_w^{mr'}} = - \frac{(\phi_w^{mr'})^{\rho})}{\phi_w^{mr'} \left((\phi_w^{mr'})^{\rho} - (\phi_m^{wr'} G)^{\rho} \right)} - \frac{(1 - \phi_w^{mr'})^{\rho}}{(1 - \phi_w^{mr'}) \left((1 - \phi_w^{mr'})^{\rho} + ((1 - \phi_w^{mr'})G)^{\rho} \right)} < 0 \tag{23}$$

$$\frac{\partial F_2}{\partial \phi_m^{wr'}} = \frac{2(\phi_w^{mr'})^{\rho} + (\phi_m^{wr'} G)^{\rho}}{\phi_m^{wr'} \left((\phi_w^{mr'})^{\rho} + (\phi_m^{wr'} G)^{\rho} \right)} + \frac{2(1 - \phi_w^{mr'})^{\rho} + ((1 - \phi_m^{wr'})G)^{\rho}}{(1 - \phi_m^{wr'}) \left((1 - \phi_w^{mr'})^{\rho} + ((1 - \phi_w^{mr'})G)^{\rho} \right)} > 0 \tag{24}$$

Appealing to Cramer's rule,

$$\frac{\partial \phi_w^{mr'}}{\partial G} = \frac{\frac{\partial F_1}{\partial G}\frac{\partial F_2}{\partial \phi_m^{wr'}} - \frac{\partial F_1}{\partial \phi_m^{wr'}}\frac{\partial F_2}{\partial G}}{\frac{\partial F_1}{\partial \phi_w^{mr'}}\frac{\partial F_2}{\partial \phi_m^{wr'}} - \frac{\partial F_1}{\partial \phi_m^{wr'}}\frac{\partial F_2}{\partial \phi_w^{mr'}}} \tag{25}$$

$$\frac{\partial \phi_w^{mr'}}{\partial G} = \frac{\frac{\partial F_2}{\partial G}\frac{\partial F_1}{\partial \phi_w^{mr'}} - \frac{\partial F_2}{\partial \phi_w^{mr'}}\frac{\partial F_1}{\partial G}}{\frac{\partial F_1}{\partial \phi_w^{mr'}}\frac{\partial F_2}{\partial \phi_m^{wr'}} - \frac{\partial F_1}{\partial \phi_m^{wr'}}\frac{\partial F_2}{\partial \phi_w^{mr'}}} \tag{26}$$

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

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

Appendix B

In this appendix we discuss four issues with the data. 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, and iv) aggregating the in-home sample to the school.

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:

$$P(MatchOut|m,s) = \frac{\exp\left(\sum_g I(G_m = g)\gamma_g + \sum_r I(R_m = r)\gamma_r + \gamma_s\right)}{1+\exp\left(\sum_g I(G_m = g)\gamma_g + \sum_r I(R_m = r)\gamma_r + \gamma_s\right)}. \tag{26}$$

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:

$$\begin{aligned} P(PastSex|m, s, MatchOut = 0) &= \frac{\exp\left(\sum_g I(G_m = g)\theta_g + \sum_r I(R_m = r)\theta_r + \theta_s\right)}{1 + \exp\left(\sum_g I(G_m = g)\theta_g + \sum_r I(R_m = r)\theta_r + \theta_s\right)} \\ &= \pi^s_{m1} \end{aligned} \quad (26)$$

again using grade, race and school fixed effects. These predicted conditional probabilities are used as weights to integrate out the number of men searching in each market.

The next data issue concerns missing reports on partner characteristics. Because of the likelihood approach it is straightforward to take what information we do observe from reports on the partner characteristics and terms and form likelihood terms which integrate out over the missing data, and enter 4.

Finally, since we only observe matches for the (individual) *in-home* sample, we must take into account the fraction of the (aggregate) *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. We estimate a first-stage logit on matching outside the school separately for men and women, incorporating grade, race and school fixed effects. Predicted probabilities are then used to adjust the aggregate number of individuals of each type searching at each school.

Table 1: Means by Gender[a]

	Men	Women
Currently		
Matched(sex or relationship)	0.337	0.315
In a Relationship	0.315	0.295
Having Sex	0.185	0.159
Prior sex	0.321	0.237
Current Sex | Race		
White	0.189	0.166
Black	0.240	0.190
Hispanic	0.167	0.122
Other	0.074	0.099
Prior Sex | Race		
White	0.277	0.228
Black	0.527	0.314
Hispanic	0.365	0.189
Other	0.191	0.136
N	3,687	3,418

[a]Sample includes only those searching in-school, under the assumption that $P^{out}_{match} = 1$. Current is defined as ongoing at the time of the in-home survey. Relationship means holding hands and kissing; sex refers to sexual intercourse.

Table 2: Stated Preferences by Gender[a]

Prefer:	Men	Women
Relationship	0.961	0.971
Sex	0.607	0.360
Sex, By Race :		
White	0.583	0.349
Black	0.738	0.404
Hispanic	0.635	0.381
Sex, By Grade :		
9th	0.487	0.254
10th	0.583	0.361
11th	0.672	0.397
12th	0.723	0.480
N	3,687	3,418

[a] 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: Conditional Means of Sex Participation[a]

Observed:	Women	Men	Difference
P(Sex|Want sex=1)	0.340	0.287	-0.053***
N	1172	2127	
P(Sex|Want sex=1,Matched)	0.738	0.682	-0.057***
N	539	895	
P(Sex|Want sex=0,Matched)	0.272	0.201	-0.071**
N	344	537	
With No Prior Sex			
P(Sex|Want sex=1,Matched)	0.579	0.458	-0.121***
N	280	389	
P(Sex|Want sex=0,Matched)	0.175	0.122	-0.053*
N	279	441	
With Prior Sex			
P(Sex|Want sex=1,Matched)	0.911	0.854	-0.057
N	259	506	
P(Sex|Want sex=0,Matched)	0.719	0.538	-0.180**
N	65	96	

[a]*,**,*** denote significance at the 5,1, and 0.01% levels respectively. Matched is defined has having either a relationship or sex in-school. Sample includes only in-school matches.

Table 4: Cross-Grade Matching Distribution.[a]

Female Grade	Male Grade 9th	10th	11th	12th	Total	Fraction of Sample
9	0.084	0.062	0.061	0.033	0.258	0.301
10	0.028	0.092	0.079	0.077	0.302	0.279
11	0.017	0.037	0.080	0.078	0.240	0.234
12	0.005	0.016	0.046	0.116	0.201	0.186
Total	0.138	0.219	0.302	0.341	1.000	
Fraction of Sample	0.276	0.280	0.246	0.198		1.000

[a]Distribution from 1077 within-school matches with valid partner information. Fraction of sample refers to sample of in-school searchers under the assumption that $P^{out}_{match} = 1$.

Table 5: Cross-Race Matching Distribution.[a]

Female Race	Male Race White	Black	Hispanic	Other	Total	Fraction of Sample
White	0.518	0.009	0.041	0.007	0.630	0.577
Black	0.006	0.157	0.008	0.001	0.194	0.199
Hispanic	0.026	0.007	0.071	0.001	0.122	0.144
Other	0.010	0.006	0.008	0.019	0.054	0.800
Frac of Matches	0.560	0.180	0.129	0.029	1.000	
Fraction of Sample	0.611	0.162	0.146	0.081		1.000

[a]Distribution from 1077 within-school matches with valid partner information. Fraction of sample refers to sample of in-school searchers under the assumption that $P^{out}_{match} = 1$.

Table 6: Variation in Gender Ratio[a]

	Percentile		
Ratio of Females to Males:	.25	.50	.75
Total	0.794	0.898	1.002
White	0.772	0.885	0.994
9th	0.780	0.917	1.077
10th	0.769	0.914	1.095
11th	0.623	0.791	0.923
12th	0.613	0.813	0.957
Black	0.729	0.930	1.096
9th	0.481	0.891	1.178
10th	0.330	0.873	1.232
11th	0.352	0.797	1.145
12th	0.102	0.701	1.013

	Fraction Female	
Overall Fraction Female	< 25th	> 75th
P(Sex|Match)	0.513	0.530
Same-Race Fraction Female		
P(Sex|Match)	0.491	0.552
Fraction Female of Partner's Race-Grade		
P(Sex|Match)	0.496	0.564

[a]Based on a sample of 73 schools. Gender ratios are calculated using only those searching within the school. Aggregate gender ratio refers to the ratio of searching females to searching males.

Table 7: Reduced Form Probability of Sex Conditional on Matching[a]

	P(Sex|Match)	
	(i)	(ii)
F/M Ratio in Partners' Grade-Race	0.089**	0.107**
	(0.044)	(0.051)
Prior Sex	0.422***	0.434***
	(0.023)	(0.107)
Grade	0.069***	0.063
	(0.014)	(0.044)
Partner Grade	0.027**	0.022
	(0.014)	(0.025)
School Characteristics	Yes	No
School Fixed Effects	No	Yes
N	893	893

[a]Coefficients 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[a]

Matching Parameters	Estimates	Standard Errors
ρ	-0.312	(0.090)
A	0.415	(0.002)
σ^{-1}	0.168	(0.003)
Sex Utility		
Male $\times$ Sex (α_7)	-3.788	(1.978)
Female $\times$ Sex (α_9)	-17.662	(1.187)
Past-Sex $\times$ Sex (α_8)	15.276	(0.973)
Male Grade $\times$ Sex (α_{12})	1.506	(0.475)
Female Grade $\times$ Sex (α_{13})	3.738	(0.481)
Match Utility		
Same grade (α_1)	4.957	(0.395)
Partner Grade $\times$ Boy (α_2)	-0.668	(0.338)
Partner Grade $\times$ Girl (α_5)	4.750	(0.243)
Same Race (α_3)	10.298	(0.506)
Partner Black $\times$ Boy (α_{4b})	-0.824	(1.426)
Partner Black $\times$ Girl (α_{6b})	4.602	(0.960)
Partner Hisp $\times$ Boy (α_{4h})	-7.276	(1.464)
Partner Hisp $\times$ Girl (α_{6h})	-3.770	(1.404)
Partner Other $\times$ Boy (α_{4o})	-10.768	(1.773)
Partner Other $\times$ Girl (α_{6o})	-7.837	(0.881)
$-\log(\ell)$	4541.72	

[a]Estimates are from sample of in-school searchers 3449 feamles in 1083 two-sided matches. Standard errors are in parentheses.

Table 9: Stated vs. Predicted Preferences for Sex[a]

	Stated		Model
Group:	Conditional	Unconditional	Predicted
Male	0.607	0.583	0.618
Female	0.360	0.346	0.316
Male			
White	0.583	0.563	0.603
Black	0.738	0.703	0.705
Hispanic	0.635	0.625	0.631
9th Grade	0.487	0.463	0.518
10th Grade	0.583	0.559	0.595
11th Grade	0.672	0.650	0.677
12th Grade	0.723	0.700	0.743
Female			
White	0.349	0.339	0.318
Black	0.404	0.379	0.355
Hispanic	0.381	0.370	0.299
9th Grade	0.254	0.242	0.153
10th Grade	0.361	0.350	0.277
11th Grade	0.397	0.380	0.419
12th Grade	0.480	0.465	0.542

[a]Subjective preference means come from sample of both men (N=3689) and women (N=3472). The conditional column includes only respondents who wanted to match, the unconditional column includes all respondents in our sample. 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]

Female Grade and Relationship:	Male Grade 9 P_m^w	9 P_w^m	10 P_m^w	10 P_w^m	11 P_m^w	11 P_w^m	12 P_m^w	12 P_w^m
9								
sex	0.07	1.00	0.13	0.95	0.22	0.72	0.33	0.51
no sex	0.23	0.69	0.38	0.45	0.63	0.26	0.95	0.14
10								
sex	0.11	1.00	0.18	0.81	0.28	0.59	0.42	0.41
no sex	0.20	0.75	0.34	0.50	0.58	0.29	0.88	0.15
11								
sex	0.12	1.00	0.19	0.79	0.29	0.57	0.44	0.39
no sex	0.15	0.90	0.26	0.62	0.46	0.37	0.73	0.21
12								
sex	0.11	1.00	0.18	0.81	0.28	0.59	0.42	0.41
no sex	0.10	1.00	0.19	0.77	0.35	0.48	0.59	0.28

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

Table 11: Varying First-Stage Assumptions[a]

	% Never Matched Removed 0		25		50	
	P_{match}^{out}		P_{match}^{out}		P_{match}^{out}	
Mean ϕ_{sex}	1	0.66	1	0.66	1	0.66
Male, No Equilibrium	0.618	0.628	0.568	0.579	0.485	0.498
Female, No Equilibrium	0.316	0.302	0.362	0.344	0.425	0.395
Mean Stated Preference						
Male	0.583	0.587	0.608	0.600	0.630	0.621
Female	0.346	0.352	0.370	0.368	0.387	0.384

[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.

Table 12: Counterfactual Probabilities of Matching: Blacks in Aggregate School[a]

	Male Grade							
Female Grade	9		10		11		12	
and Relationship:	P_m^w	P_w^m	P_m^w	P_w^m	P_m^w	P_w^m	P_m^w	P_w^m
9								
sex baseline	0.11	1.00	0.19	0.78	0.31	0.55	0.47	0.36
no sex baseline	0.33	0.51	0.57	0.29	0.90	0.15	1.00	0.05
sex CF Only GR	0.10	1.00	0.17	0.83	0.28	0.59	0.43	0.40
no sex CF Only GR	0.31	0.54	0.52	0.32	0.85	0.17	1.00	0.06
sex CF Full	0.10	1.00	0.17	0.83	0.28	0.60	0.42	0.41
no sex CF Full	0.27	0.60	0.46	0.37	0.75	0.20	1.00	0.09
10								
sex baseline	0.15	0.90	0.24	0.66	0.38	0.45	0.57	0.29
no sex baseline	0.29	0.58	0.50	0.34	0.82	0.17	1.00	0.07
sex CF Only GR	0.13	0.95	0.21	0.72	0.34	0.50	0.52	0.33
no sex CF Only GR	0.26	0.62	0.45	0.38	0.76	0.20	1.00	0.08
sex CF Full	0.14	0.94	0.22	0.71	0.34	0.50	0.51	0.33
no sex CF Full	0.24	0.67	0.40	0.43	0.67	0.23	1.00	0.12
11								
sex baseline	0.15	0.90	0.24	0.66	0.38	0.45	0.57	0.29
no sex baseline	0.21	0.74	0.38	0.45	0.65	0.25	1.00	0.12
sex CF Only GR	0.13	0.96	0.21	0.73	0.34	0.50	0.52	0.33
no sex CF Only GR	0.18	0.79	0.34	0.50	0.59	0.28	0.93	0.14
sex CF Full	0.14	0.93	0.22	0.71	0.34	0.50	0.51	0.33
no sex CF Full	0.17	0.83	0.30	0.55	0.53	0.32	0.83	0.17
12								
sex baseline	0.14	0.95	0.22	0.71	0.35	0.49	0.53	0.32
no sex baseline	0.14	0.92	0.28	0.59	0.50	0.34	0.81	0.18
sex CF Only GR	0.12	1.00	0.19	0.77	0.31	0.54	0.48	0.36
no sex CF Only GR	0.13	0.98	0.25	0.65	0.45	0.38	0.75	0.20
sex CF Full	0.13	0.97	0.20	0.74	0.32	0.53	0.49	0.35
no sex CF Full	0.12	1.00	0.22	0.71	0.40	0.43	0.67	0.24

[a]CF 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 individuals' grade and possible partner grade. P_m^w is the probability of matching for a man looking for a woman, P_w^m is the probability of matching for a woman looking for a man.

Table 13: Racial Gap in the Probability of Sex Conditional Matching Under Various Market Conditions

	Same Type P(Sex|Match)[a]			
Aggregate School	9	10	11	12
White	0.210	0.393	0.537	0.648
Black	0.312	0.497	0.644	0.741
Difference	-0.102	-0.105	-0.106	-0.093
Changing: Only Gender Ratios				
White	0.209	0.393	0.537	0.648
Black	0.298	0.491	0.638	0.738
Difference	-0.088	-0.099	-0.101	-0.090
Gender Ratios and Male Past Sex				
White	0.209	0.393	0.538	0.648
Black	0.274	0.431	0.573	0.672
Difference	-0.065	-0.038	-0.035	-0.024
Gender Ratios and All Past Sex				
White	0.209	0.393	0.538	0.648
Black	0.239	0.401	0.544	0.652
Difference	-0.030	-0.008	-0.006	-0.004

[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.