STRATEGIC INTERACTION IN THE SEX MARKET

JOHN MORROW† AND YOAV SIVAN††

July 31, 2007

ABSTRACT. There have been few attempts to empirically explain the pursuit of short term relationships and sex in a formal context. Previous work has lamented the paucity of empirical studies which utilize incentive driven behavior to draw conclusions and recommend policy. We develop a model of social network formation through sexual matching, provide an empirical approach derived from the model and apply it to a population of high interest. Specifically, we apply the approach to a population of sexually active men who have sex with men (MSM) in a large metropolitan area and derive qualitative conclusions regarding how individuals behave in the marketplace for sex.

† Department of Agricultural and Applied Economics, UW-Madison, morrow1@wisc.edu.
††Columbia School of Journalism, yoav76@gmail.com.

Acknowledgement. We would like to thank the following for invaluable comments and discussions Dan Ariely, Michael Carter, Swati Dhingra, Steven Durlauf, Shachar Kariv, Rebecca Lessem, Ching-Yang Lin, Jack Porter, Laura Schechter and Ken West as well as other students and faculty who have offered several helpful comments. All errors are our own.

1. Introduction

"Sex without love is an empty experience, but as empty experiences go, its one of the best."

- Woody Allen

We develop a model of social interaction between self-interested agents in what can be termed as the "sexual marketplace." (Laumann, 2004) In particular, we investigate patterns of sexual matches between men who have sex with men (MSM). MSM engage in multiple short term relationships, the pursuit of which can be considered as repeated play of one shot matching games among similarly motivated agents. The concept of such "sex marketplaces" has been primarily pursued in a sociological or epidemiological context (Laumann, 1994; Laumann and Michael, 2001; Laumann, 2004). Little work has been done in economics to model sexual behavior for its own sake[1]. In fact, Phillipson, et al. (1994) claim that AIDS policy is ineffective due to the fact that policy does not take into account how individual behavior responds to incentives (Philipson, Posner, and Wright, 1994). Although more recent work has taken this issue to task, to our knowledge little work has been done in a game theoretic context that emphasizes probabilistic short term sexual behavior[2] (Kremer, 1996). The work closest to our approach is a model involving individual choice with regard to HIV testing(Caplin and Eliaz, 2003). Our emphasis is to model sexual partner choice and consequent outcomes.

Due to the high frequency of short term relationships that MSM often engage in, patterns of frequency across individuals are more stark than in heterosexual relationships. MSM are thus a population well suited to study the determinants of *sexual* matching in comparison to matching that very often results in long term relationships. To put it bluntly,

"Women need a reason to have sex. Men just need a place."

- Billy Crystal

Modelling the dynamics of the sex market can help predict the frequency and distribution of sexual partners which occur as outcomes of the sex market. Such dynamics therefore have sociological and epidemiological implications. We ask the following questions:

(1) Based on observable characteristics, which partners will an individual pursue? In other words, how do people choose sexual partners under the resource constraint of the pool of prospective partners?
(2) Given empirical characteristics (e.g. physical appearance), in what ways do we observe agents trading off between the attractiveness of partners and the possibility of remaining single? Specifically, do real world relationships cluster around similar characteristics (is the sex market assortative)?

[1]One notable exception is Mialon (2003) (Mialon, 2003).

[2]Outside of economics, some recent work has incorporated probabilistic methods into Gale and Shapley's classic model(Immorlica and Mahdian, 2005).

In regard to the first question, we provide a formal model which allows for multiple optimal behaviors, ranging from agents trying to date only the most attractive members of the population to "playing it safe" and only dating those most similar in attractiveness to themselves. In regard to the second question, we use our model along with data on sexual behavior to estimate a key model parameter which gives a qualitative explanation to how the sex market functions. We find that the sex market is fairly assortative and estimate how often and under what conditions it is assortative. In doing so, we resolve two competing claims regarding what drives behavior in the sex market, in the sense of empirically estimating conditions under which each occurs.

This paper is organized as follows: A formal model with emphasis on examples is presented, along with a few key results which drive the model. Next is a specification of the statistical test following from the model and a description and discussion of the data used. Then results are presented and interpreted in model context, followed by a discussion of two competing claims and areas for further work. Finally, a detailed appendix[3] provides further model examples and results for the interested reader.

2. Model Synopsis

In AgentLand there is a population P full of lonely agents with different attractiveness types $\theta_i > 0$. Agents can end their loneliness by matching with each other, and agent i receives a utility $u(\theta)$ from matching with type θ[4]. Thus, when two types θ_i and θ_j match with one another, agent i receives a payoff of $u(\theta_j)$, and agent j receives a payoff of $u(\theta_i)$. "Self matching" is not allowed. In keeping with θ being a measure of attractiveness, we assume that u is strictly increasing in θ. Accordingly, everyone would like to match with the highest type possible. By convention we label the types θ_i in increasing order, i.e. $\theta_1 < \theta_2 < \ldots$. At the end of the game, an agent i may remain lonely (unmatched) in which case we say they are matched to the "null agent" $\theta_0 = 0$ and receive a payoff of $u(\theta_0) = 0$. Notice this assumes that "any match is better than no match," an assumption most suited to short term rather than long term behavior[5].

Unfortunately for the lonely agents, in order to match, at night they must go to one of the local bars in AgentLand, each of which can only hold N agents.

2.1. An evening at the AgentBar.

Once a bar in AgentLand fills up with N agents, agents meet each other and each agent's type θ_i becomes public knowledge. During the course of the evening, each agent i may proposition a different agent j, and we agent i's choice as $A_i \in \{1, 2, \ldots, N\}$. Agents choose A_i to maximize expected utility as developed below. When agent i propositions agent j, we call the proposition a *contact* from i to j. Immediately before closing time, we can summarize the various contacts across agents in the *contact network* depicted in Figure 2.1 for $N = 4$.

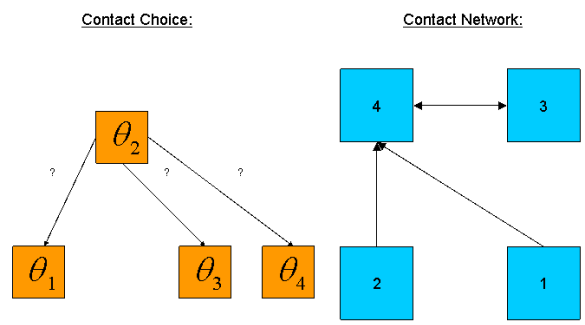
Contacts are *successful* or *unsuccessful* with probabilities p and $1 - p$, respectively and the *success* of a contact is not known to any agent until the bar closes. In other words,

[3]To be interpreted as "available upon request," but included with this paper for completeness.

[4]Here we have assumed that all agents have the same preferences, but it may not be too difficult to find data regarding preferences in sexual roles which have some explanatory power, see for instance (Yee, 2002).

[5]Recent research suggests that cues regarding short and long term intent are visually observable, e.g. (Warner, 2006).

FIGURE 2.1. A Contact Network.



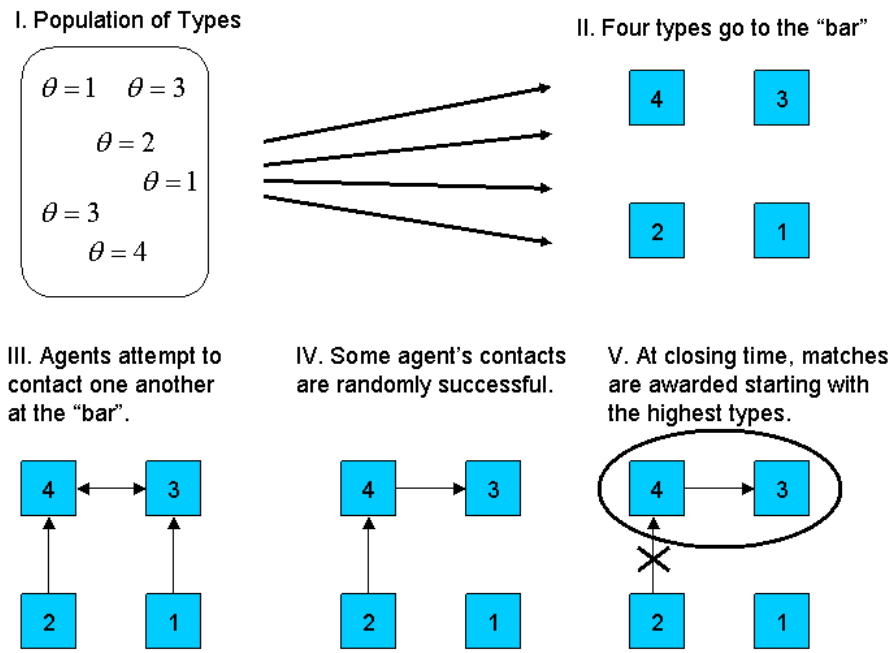
p is a "noise parameter" determining the success or failure of any particular proposal to match and the closer p is to one, the more likely any particular proposal is successful. In this paper we take p as exogeneously given for two reasons. First, we will in fact estimate p for a particular population below and our dataset is not sufficiently rich to reasonably estimate an endogeneous p. Second, by considering p as exogeneous for *particular venues* we can predict different behaviors in different venues. For example we might consider internet dating with a high p (multiple body photographs, sober individuals, detailed conversations) versus the "darkrooms" found in gay bars with low p (dim lights to hide characteristics, alcohol and drug use, very little conversation).

2.2. **Closing time.** At closing time, agents have exhausted all of their strategic actions and nature takes its course. First, *unsuccessful* contacts are removed from the contact network. Second, starting with the highest type agent, that agent is awarded a match with their highest *successful* contact (if one exists) and the matched pair is removed from the network. Third, the second stage is repeated for the next highest type agent remaining in the network. This process is formalized in Algorithm 1. We summarize all steps of the game in Figure 2.2.

2.3. **Beliefs and Equilibrium.** Above we have specified how matches occur for any particular contacts A_i that agents choose, as well as the payoffs obtained by matching (or remaining unmatched). In order to close the model, we need to specify how agents choose A_i. First, we assume that agent i has subjective beliefs μ_i regarding the choices of all agents. Since each agent chooses $A_i \in \{1, \ldots N\}$, we may define a belief μ_i as a probability density over $\times_{i=1}^{N} A_i = \{1, \ldots, N\}^N$. Since each vector in $\times_{i=1}^{N} A_i$ is nothing but the list of contacts made by agents and each such vector corresponds to a unique Contact Network, we may think of a belief μ_i as a subjective probability density over Contact Networks. For simplicity we will only consider beliefs which place probability one on a particular $(A_1, \ldots A_N) \in \times_{i=1}^{N} A_i$.

A belief μ_i paired with a choice A_i generates expected utility $U(A_i, \mu_i)$ which we develop below. The appropriate equilbrium concept is choices $A^* \equiv (A_1^*, \ldots, A_N^*)$ and

FIGURE 2.2. Steps of the Game for $N = 4,\ k = 1$



Algorithm 1 Resolving the Contact Network.

(1) Make a graph with each vertex representing an agent which contains no edges.
(2) Look at each agent's choice A_i, and independently with probability p each contact is represented as a directed edge going from the agent making the contact to who they contact. For example, a total of k edges are drawn with probability $\binom{N}{k} p^k (1-p)^{N-k}$.
(3) Start with the highest numbered agent, say i, and find the edge which is listed highest in their contact list, say E. Now one of three things can happen:
  (a) If the agent has no outgoing edges, they are free for subsequent matches.
  (b) If E has an outgoing edge to an unmatched agent j, i and j are matched.
  (c) If E has an outgoing edge to a matched agent j, agent i goes home alone and is considered matched in evaluating subsequent agents.
(4) Repeat Step 3 until all agents have been evaluated.

beliefs $\mu \equiv (\mu_1, \ldots, \mu_N)$ such that choices are optimal given beliefs. Formally, (A^*, μ) constitute an equilibrium when for each agent i we have

$$U(A_i^*, \mu_i) \quad = \quad \max_{j \in \{1,\ldots,N\}} U(j, \mu_i)$$

In order to pare down the number of equilibria in this game[6] we propose the refinement of a Rational Expectations (RE) equilibrium, which requires that agents beliefs match up with other agent's chosen actions. Formally,

Definition. (RE Equilibrium) An equilibrium (A^*, μ) is a Rational Expectations equilibrium if and only if $\mu_i(A_1^*, \ldots, A_N^*) = 1$ for each agent i.

It is shown below that a Rational Expectations equilbrium always exists. Hereafter we consider only RE equilibria.

2.4. **Expected Payoffs.** Having specified the action A_i open to each agent and his beliefs μ_i, we can define the agent's expected payoff for each such pair (A_i, μ_i). From each agent's perspective, every pair (A_i, μ_i) induces an N by N matrix $C(A_i, \mu_i)$ which contains the probabilities of various agents matching at closing time. Specifically, $C(A_i, \mu_i)_{j,k}$ is the probability that agent j matches with agent k *as a result of* j contacting k.

In order to make this concrete, refer to the Contact Network in Figure 2.1 with agent types 1,2,3, and 4 so $\theta_1 = 1$, $\theta_2 = 2$, etc. Suppose that the agent with type 1 has beliefs which are consistent with the lines drawn and chooses $A_1 = 4$, as in the Figure. Then we have that (applying the clearing rule specified above)[7]

$$C(A_1 = 4, \mu_1) = \begin{bmatrix} 0 & 0 & 0 & p(1-p)^3 \\ 0 & 0 & 0 & p(1-p)^2 \\ 0 & 0 & 0 & p(1-p) \\ 0 & 0 & p & 0 \end{bmatrix}$$

Still referring to Figure 2.1 suppose that the type 1 agent instead contacted agent 2. Then the probability of agent 1 matching with agent 2 *resulting from* agent 1 contacting agent 2 would be (again, recall the clearing rule):

$$\Pr(\text{contact from 1 to 2 successful}) \cdot \Pr(\text{2's contact was unsuccessful and 2 is unmatched}) = p \cdot (1-p)$$

so that the new matrix is

$$C(A_1 = 2, \mu_1) = \begin{bmatrix} 0 & p(1-p) & 0 & 0 \\ 0 & 0 & 0 & p(1-p)^2 \\ 0 & 0 & 0 & p(1-p) \\ 0 & 0 & p & 0 \end{bmatrix}$$

Now as far the agent is concerned, he only cares about the net probability of matching with another agent and the type of that agent, regardless of who made contact. Accordingly, if we define $M(A_i, \mu_i)$ by

$$M(A_i, \mu_i) \equiv C(A_i, \mu_i) + C(A_i, \mu_i)^\top$$

we arrive at a matrix which contains the net probability of i and j matching in its (i^{th}, j^{th}) place. Finally, defining a vector of utilities $\tilde{u}$ by

$$\tilde{u} \equiv (u(\theta_1), \ldots, u(\theta_N))$$

[6]Clearly we can have multiple equilibria from "crazy" beliefs. E.g. say the least attractive individual believes everyone is contacting him which results in suboptimal behavior.

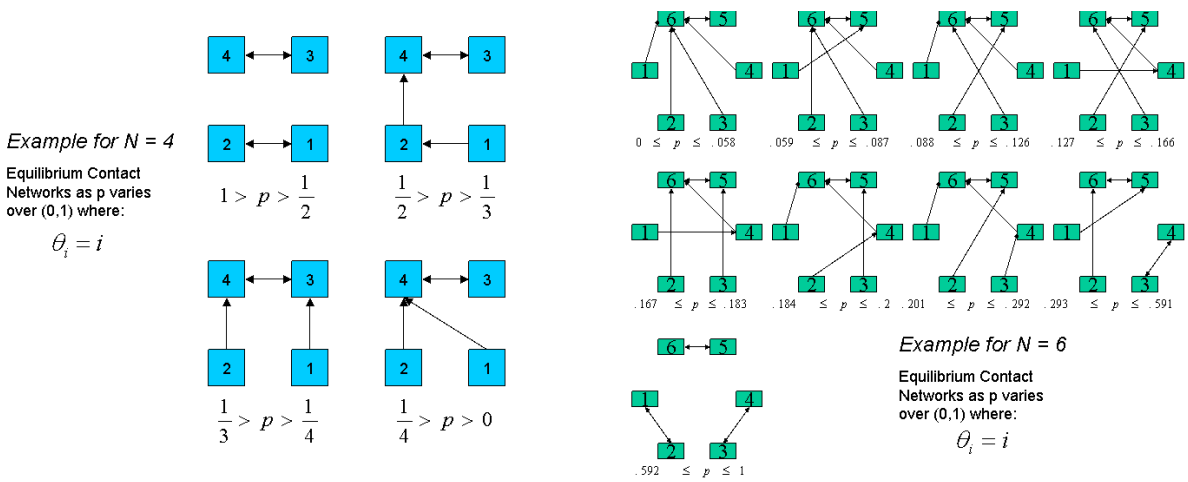
[7]In examining the matrix, it helps to start with the bottom row.

we see that conditional on the action A_i and beliefs μ_i, the expected payoff to agent i is the ith element of the vector $\tilde{u}M(A_i, \mu_i)$. We therefore define agent i's expected payoff as (where e_i is the ith standard basis vector)

$$U(A_i, \mu_i) \equiv \tilde{u}M(A_i, \mu_i)e_i \tag{2.1}$$

2.5. **Model Examples.** To illustrate the model more fully[8], we present many different equilibrium outcomes for different values of p in Figure 2.3. Each box in the Figure represents an agent of type $\theta_i = i$ and a directed edge from θ_i to θ_j indicates that θ_i has chosen to contact θ_j, or rather $A_i = j$. Except at possibly the boundary values of p, each equilibrium illustrated is unique for the range of p stated.

FIGURE 2.3. Specific Examples of Equilibria Across p.



(A) Four Agents (B) Six Agents

As above, each contact network induces a distribution of partners for each agent, given by $M(A_i, \mu_i)$. Since we are assuming Rational Expectations, $M(A_i, \mu_i)$ is the same for each agent i. For the six agent case above with $p = .5$ the relevant distribution is

$$M(A_i, \mu_i) = \begin{bmatrix} 0 & 0 & 0 & 0 & .125 & 0 \\ 0 & 0 & 0 & 0 & 0 & .125 \\ 0 & 0 & 0 & .75 & 0 & 0 \\ 0 & 0 & .75 & 0 & 0 & 0 \\ .125 & 0 & 0 & 0 & 0 & 0.75 \\ 0 & .125 & 0 & 0 & 0.75 & 0 \end{bmatrix}$$

For instance, $M(A_i, \mu_i)_{3,4} = .75$ indicates that the probability of types 3 and 4 being matched is .75, which follows from the following calculations:

Agent $i = 3 \longrightarrow$Agent $i = 4$:

$$\Pr((3,4) \text{ Match from A3 contacting A4}) = .5$$

[8]Details illustrating the analytical calculations for the $N = 4$ case of Figure 2.3 can be found in the Supplemental Appendix. The equilibria for the case $N = 6$ have been numerically solved via a simulation program.

Agent $i = 4 \longrightarrow$Agent $i = 3$:

$\Pr((3,4)\text{ Match from A4 contacting A3}|\text{A4 is unmatched})\Pr(\text{A4 is unmatched}) = .25$

The raw probabilities of agents matching with anyone whatsoever are then (where the ith place holds the match probability of agent i):

$$M(A_i, \mu_i) \cdot \mathbf{1} = \begin{bmatrix} .125 & .125 & .75 & .75 & .875 & .875 \end{bmatrix}^\top$$

In this case the least desirable types θ_1 and θ_2 have traded the high probability of being matched to one another in the slim hopes of being matched to higher types. The lesson is that desirable agents will seldom go unmatched and should therefore over time have more matches from a wider selection of agents. This suggests for instance that the most desirable agents are important targets for effective intervention in preventing Sexually Transmitted Infections.

3. Model Implications

In a model of sexual matching, common experience as well as intuition suggests the following features should be observed:

(1) *(Personal Standards)* In equilibrium, a given attractiveness type should never contact a type which is "below their standards," although he might match with slightly lower types when better options have been exhausted.
(2) *(Reciprocity)* When a higher type contacts a lower type, the lower type should in some way reciprocate.
(3) *(Chasing and Sorting)* When the level of "noise" is high, all agents will take a chance *chasing* the highest agent, and when noise is low will *sort* perfectly in pairs consisting of similar levels of attractiveness.

In this Section we show how each of these behavioral observations are generated by the model.

3.1. Personal Standards.

Consider the most attractive agent in the game, agent N. He is blessed by having a high type in that his choice A_N is evaluated first at closing time and therefore if successful he will be matched to agent A_N. Since A_N thereby "trumps" all other agents, the optimal choice for agent N is $A_N = N - 1$ since agent $N - 1$ is the most attractive of the other agents and therefore gives the highest payoff. Using similar logic for agent $N - 1$, he might choose $A_{N-1} = N$ or $N - 2$ but will not choose $A_{N-1} < N - 2$ in equilibrium since it must be that $A_N > N - 2$ so that agent $N - 1$ "trumps" all agents but agent N. Proceeding inductively, we conclude a formal result that says it is never optimal for an agent to lower their standards to contacting anyone who is much less attractive than themselves[9].

Proposition 1. *(Personal Standards) Let A_i^* be agent i's optimal choice in equilibrium. Then*

$$A_i^* \geq i - 1$$

In other words, an agent will never contact an agent who is more than one "level of attractiveness" below his own.

[9]We provide a more rigorous proof of this result in the Appendix.

3.2. **Reciprocity.** Our result regarding Personal Standards underscores the fact that optimal behavior takes into account both the contacts of other agents and knowledge of the "pecking order" present in the sex market. Another salient prediction of the model is the *reciprocity* exhibited by less attractive agents towards more attractive agents. The proof is more complex and is therefore in the appendix, but the rationale is as follows: When an attractive agent's best option is to contact a less attractive agent the less attractive agent has even worse options except for reciprocating. This makes reciprocation a best response for the less attractive agent. We summarize the result in Proposition 6.

Proposition 2. *(Reciprocity) In equilibrium if a more attractive agent contacts a less attractive agent, the less attractive agent reciprocates with a contact. Formally, if $A_i = j < i$ in equilibrium then $A_j = i$.*

Proof. See Appendix. □

3.3. **Chasing and Sorting.** For large p, contacts are highly successful so that when high types contact each other, there is very little chance of high types being available to low types. This results in cliques of similar agents contacting each other, and we label these phenomena "sorting equilibria." Such equilibria are illustrated in Figure 2.3 for the highest values of p. As p decreases, high types are more likely free when low types' contacts are evaluated so low types "deviate" from cliques to chase high types. For sufficiently small p, all types except the highest chase the highest type[10], and we label this a "chasing equilibrium". Again such equilibria are illustrated in Figure 2.3 for the lowest values of p. We capture these qualitative aspects of large and small p cases with Proposition 7.[11]

Proposition 3. *Define the "2-person clique" function ρ by*

$$\rho(i) = \begin{cases} i+1, & i \text{ odd} \\ i-1, & i \text{ even} \end{cases}$$

Assume N is even. Then $\exists \epsilon_c, \epsilon_s > 0$ such that in any equilibrium

(1) *All equilibria are "chasing equilibria" if $p \in (0, \epsilon_c)$, formally $A_i = N \quad \forall i < N$.*
(2) *All equilibria are "sorting equilibria" if $p \in (1-\epsilon_s, 1)$, formally $A_i = \rho(i)$.*

Proof. See Appendix.

□

An explicit example relating "chasing" and "sorting" equilibrium to the *Tragedy of the Sexual Commons* may be found in the Appendix.

[10] This also occurs if the highest type has a sufficiently large θ for any fixed p.

[11] We assume N is even for brevity. A similar statement is true for N odd.

4. Data Description

We wish to explain outcomes in the sexual marketplace as a result of rational self-interested individuals in the pursuit of sexual relationships. This necessitates a classification of individual characteristics that, through some sort of mechanism, has an explanatory effect on relationship outcomes. Following literature that specializes in MSM behavior, the primary characteristic we focus on is physical attractiveness which has been shown to have significant behavioral implications (Sergios and Cody, 1986).

We use data from sexually active MSM surveyed in the 1997 Chicago Health and Social Life Survey (CHSLS) using a suitably defined subpopulation of sexually active MSM[12]. The study is available online and contains a wealth of variables including number and characteristics of partners, satisfaction in the relationship, and individual demographics. We acknowledge that survey data is inherently flawed due to the nature of self-reporting, but it is the best option since we don't have data regarding sexual behavior which is recorded by an impartial observer.

In this section we describe the characteristics of our population, define our main variables of interest, and discuss the properties of those variables.

4.1. **Population Characteristics.** Our sample consists of 55 MSM aged 22 to 54 who were interviewed in the Chicago area. The median individual was white, 33 years old, had a median income of $30,000-40,000 and a terminal bachelor's degree. Of more direct interest, the median individual had 4-5 sexual partners over the year preceding the survey (a variable we label part_i) agreed that "sex is very important to me" and thought about sex several times per day[13].

4.2. **Primary data definitions.** In order to model matching behavior, we need to answer the question: "Who is matching with who, and how successfully?" In this paper "who" means what level of physical attractiveness one possesses. "How successfully" means "was a match good enough to match again with the same person," in other words one's relationship status with a person of given physical attractiveness.

Physical Appearance. The CHSLS data set asks each subject to rate himself as well as his last two partners on physical attractiveness relative to their age. In keeping with thinking of physical attractiveness as an agent's "type," we label these variables as θ_i^0 for subject type and θ_i^1 for each subject's most recent partner, respectively. Note that the index i refers to each subject in the population. Physical attractiveness is on a scale of 1 to 7. Casual examination of the CHSLS data indicates that more attractive individuals are overrepresented as sexual partners. A simple T-test shows that the mean partner appearance is greater than the subject appearance (at the 95% level). This suggests that the number of partners an individual has as well as his "quantity of sex consumed" should increase with the attractiveness of individual. This is due to greater representation of more versus less attractive partners in that individual's pool of partners, as well as his greater representation in others' pool of partners. If this is the case, we cannot use models which imply a constant percentage of single individuals across type, for instance, Shimer and Smith's (2000) search model which would be a clear candidate for this type of work (Shimer and Smith, 2000). We provide a formal setting which predicts what we casually observe and

[12]Details of our restrictions can be found in the Supplemental Appendix.

[13]We thank a female reviewer for pointing out that this "sounds normal for any guy."

the data illustrates: People with different levels of attractiveness consume different quantities and qualities of sex.

Since physical appearance is defined as type in our model, our results depend upon how we map reported physical appearance to the utilities of players in our game. Each game outcome is invariant to scaling all the types by a constant. Therefore, we need only be concerned with the relative rankings we assign to physical appearance. We have taken a natural approach, which is to assign each θ_i^0 the average percentile score they represent in the sample. This is a useful metric to use for comparability across studies or in combining datasets since it "automatically" makes attractiveness scores comparable if samples are random. We summarize this assignment of numerical values to physical attractiveness in Table 1.

TABLE 1. Conversion of Physical Attractiveness to Model Type.

Physical Attractiveness	% in Sample		Model Type	Value
4	20.4	$\longrightarrow$	θ_1	.102
5	50.0	$\longrightarrow$	θ_2	.454
6	24.1	$\longrightarrow$	θ_3	.824
7	5.6	$\longrightarrow$	θ_4	.972

Relationship Status. We investigate the premise that all individuals are involved in sexual relationships with the same frequency. Assume that at a given time the probability of someone being in a relationship is equal to q for all individuals so that whether someone is in a relationship is a Bernoulli random variable with parameter q. The alternative is that for some pair of physical attractiveness types θ and θ', we have $q_\theta \neq q_{\theta'}$. We have also constructed the variables SINGLE1$_i$ and SINGLE2$_i$, which capture current relationship status up to each subject's two most recent partners. SINGLE1$_i$ is equal to one if the subject did not expect to have sex again with his most recent partner. SINGLE2$_i$ is equal to one if the subject did not expect to have sex again with either of his two most recent partners[14]. We summarize the relationship rates in Table 2[15].

We conclude that less physically attractive individuals are more often single under both measures of "singledom" and have substantively higher rates of being single than more physically attractive individuals. As one would expect, it appears that better looking individuals are more often in an active sexual relationship. It is therefore desirable that any model describing the sex market should imply higher rates of sexual relationships for more attractive individuals. Our model does in fact satisfy this criterion.

For ease of reference, we summarize the data variables discussed in this section in Table 3.

[14]The first variable of choice to indicate relationship status for Male-Female couples, namely marriage, is not applicable to this subpopulation which is in general barred from legally recognized marriage. The question as to what effects the difference in legal institutions available to Male-Female and same sex couples have on relationship outcomes is interesting, but one we do not tackle here.

[15]Unsurprisingly, the unique individual with $\theta_i^0 = 3$ which we have left out of the analysis was single under both definitions.

TABLE 2. Relationship Frequency and Looks.

		SINGLE1$_i$	
Looks (θ^0)	Obs	% Single (q_{θ^0})	95% Binomial CI
4	11	82%	[48%,98%]
5	27	26%	[11%,46%]
6	13	23%	[5%,54%]
7	3	0%	[0%,63%] (1-tailed)

		SINGLE2$_i$	
Looks (θ^0)	Obs	% Single (q_{θ^0})	95% Binomial CI
4	11	73%	[39%,94%]
5	27	19%	[6.3%,38%]
6	13	15%	[1.9%,45%]
7	3	0%	[0%,63%] (1-tailed)

TABLE 3. Primary Variable Definitions.

Subject Types	θ_i^0
Most Recent Partner Type	θ_i^1
Second Most Recent Partner Type	θ_i^2
Relationship Status (Most Recent)	SINGLE1$_i$
Relationship Status (Two Most Recent)	SINGLE2$_i$

4.3. **Choice of Relevant Data Variables.** The variables most relevant to testing the model are θ_i^0, θ_i^1, and Single1$_i$. We believe that using θ_i^1 and Single1$_i$ best captures the idea of short term matching outcomes explained in the model. Expanding the estimation to include θ_i^2 and Single2$_i$ doubles the size of the sample, but this specification would be most appropriately used in a model which accounts for polyamory in a strategic setting, something beyond the scope of our present model.

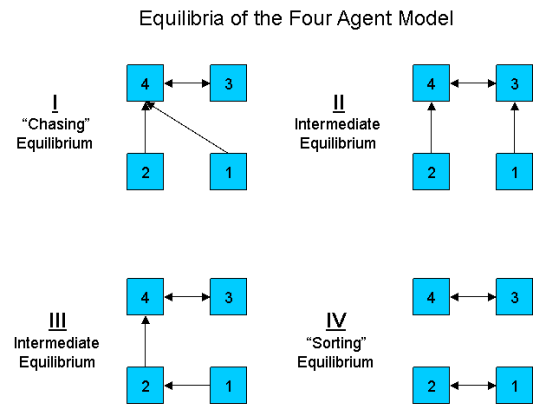
5. METHODOLOGY

Here we detail our current approach to fitting the model to data and describe the relationship of parameter estimates to qualitative features of the model.

5.1. **Choice of Estimation Parameter.** The point of our model is to provide a qualitative explanation of how the sex market functions. The model predicts a range of behavior depending on the "noise parameter" p which dictates the equilibrium that obtains whenever any group of agents meets in the sex market. It is useful to think of p as an "equilibrium parameter" which fits model equilibria to observed data. By estimating p through the process detailed below, we provide a qualitative picture of behavior within the context of the model. For estimation purposes, we focus on the 4 agent version of the model. In this case, it turns out for any 4 fixed agents there are at most 4 equilibria that occur as p

ranges over $(0,1)$[16]. Given agents with types $\{\theta_i\}$ indexed from low to high, we number these 4 equilibria from $I - IV$ and depict them below, where the indexes i represent the agent with type θ_i and an edge directed from i to j indicates i has an optimal strategy in contacting agent j.

FIGURE 5.1. Characterization of Equilibria.



Explicit calculation yields the values of p which generate each of the I-IV equilibria as summarized in Table 4. For some draws of Θ_j the II and III equilibria do not occur, which is precisely when the corresponding sets for the II and III equilibria in Table 4 are empty. For instance if $\theta_1 = \theta_2 < \theta_3 = \theta_4$ we see that Equilibrium III can never exisCalculations using Table 4 will show that as p increases in the 4 person game, we move through equilibria in the order

$$I \longrightarrow II \longrightarrow III \longrightarrow IV$$

TABLE 4. Location of equilibria.

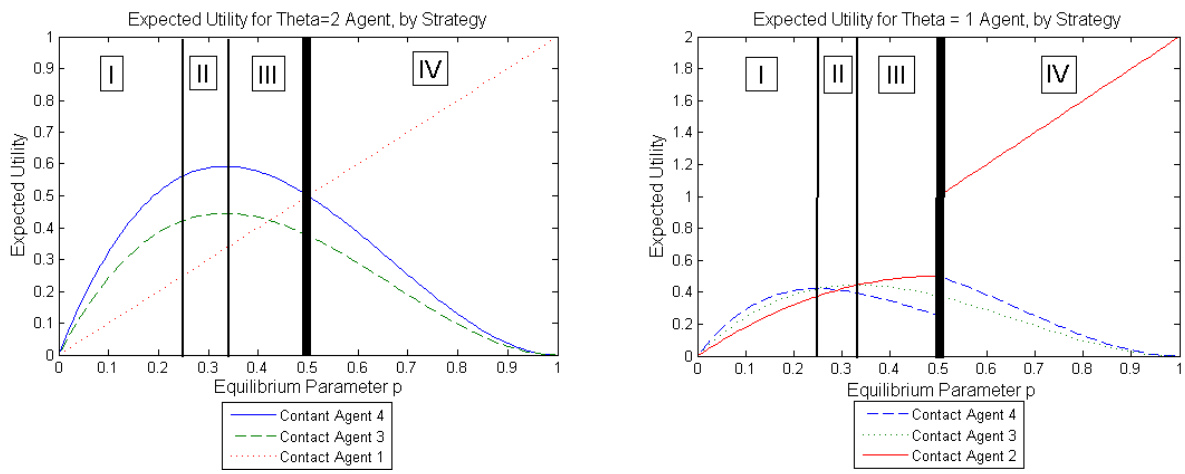
Equilibrium	$(1-p)^2$ contained in:
I	$(\frac{\theta_2}{\theta_4}, 1) \bigcap (\frac{\theta_3^2}{\theta_4^2}, 1)$
II	$(\frac{\theta_1}{\theta_4}, 1) \bigcap (\frac{\theta_2^2}{\theta_3^2}, 1) \bigcap (0, \frac{\theta_3^2}{\theta_4^2})$
III	$(\frac{\theta_1}{\theta_4}, 1) \bigcap (0, \frac{\theta_2}{\theta_4}) \bigcap (0, \frac{\theta_2^2}{\theta_3^2})$
IV	$(0, \frac{\theta_1}{\theta_4})$

After any particular draw Θ_j, knowing p determines the equilibrium in the game. Notice that the highest (θ_4) type and second highest (θ_3) type always contact one another, and that therefore the equilibrium outcome is a function of the choices of the lowest and second lowest types (θ_1, θ_2). For the case $\theta_i = i$ we illustrate the payoffs to Agents 1 and

[16]The number of equilibria ranges from 2 to 4 in the 4 agent case. For larger numbers of agents, the number of equilibria grows quite quickly which complicates the qualitative interpretation of any particular equilibrium.

2 across p in Figure 5.1. The Roman numerals correspond to the outcome equilibria as numbered in Figure 5.2. Notice the thick vertical line that occurs where Agent 2 changes strategies in equilibrium. This abruptly alters Agent 1's payoff structure due to a change in the structure of the contact network.

FIGURE 5.2. Agents' Optimal Choices and Equilibria.



(A) (B)

We assume that the data represents the outcome of independently played games consisting of random collections of agents $\Theta_j = \{\theta_i^j\}$ where for each j, θ_i are drawn with replacement from the empirical distribution of subjects. By estimating p we arrive at a behavioral prediction for any particular agents Θ_j who meet to play the game[17]. We are looking for the p that across all random draws of agents from the empirical population best fits some statistic of interest. We attempt to capture what different values of p mean for (hypothetical) equilibrium outcomes pictorially in Figure 5.3.

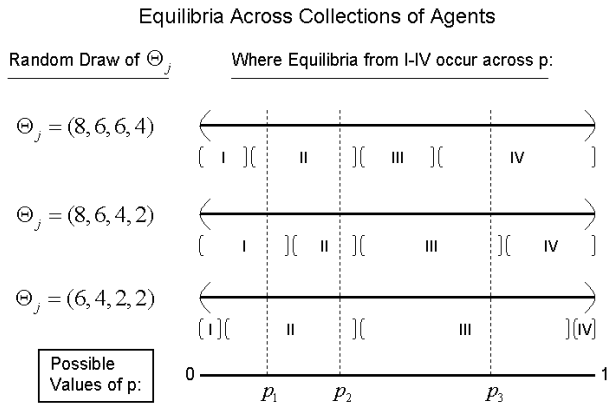
5.2. **Estimation Procedure.** We now need to specify some observable prediction of the model with which to estimate p. In this section we first convert the equilibria outcomes into a matrix form. Second, we reduce the matrix form to something which can be reasonably estimated given the size of our dataset and provide an explicit estimation procedure for p.

5.3. **Observable Predictions.** In estimating p, we need to match theoretical predictions from the model with observed data. Given p and a given set of N agents Θ we assume the agents play the game specified in the model with agents types in Θ. We also assume that the game's outcome is the unique (up to some tie-breaking) RE equilibrium. As detailed above, for any fixed p this results in a matrix M of predicted match frequencies.

Since all we observe are actual match outcomes across types, we are interested in what outcomes the model predicts for each type. Necessarily we need to convert the "one-shot" model predictions for each subset of N agents into what should be observed when different collections of N agents play the game. We take the following semi-parametric

[17]The explicit details of when and where each equilibrium occurs for a particular Θ_j draw in the 4 person game may be found in Appendix III.

FIGURE 5.3. Location of Equilibria Over the Population.



approach, which assumes that the observed distribution of types is a good representation of the true population:

Let T denote the population of types observed in the sample (how such types are determined is detailed below). We will denote by $|T|$ the number of unique types in the population and label each different one by t_i. Let $g(\theta)$ denote the empirical distribution of subject types in T and define the probability of some N agents Θ playing the game together as

$$f(\Theta) = \prod_{\theta_i \in \Theta}^{N} g(\theta_i)$$

Thus f is the distribution of N types randomly drawn from the sample with replacement. For each fixed $\Theta$and p create the matrix of equilibrium match frequencies $M(\Theta, p)$.

Now take $M(\Theta, p)$ and create the "match frequency matrix for all types in T", which we will call $\mathcal{M}(\Theta_j, p)$. Specifically, this is a $|T| \times |T|$ matrix containing the average probability of an agent of type t_i matching with type t_j in its ith, j^{th} spot.[18] Now take the expectation of $\mathcal{M}(\Theta_j, p)$ over f to arrive at the "expected match frequency matrix"

$$\tilde{\mathcal{M}}(p) \equiv \sum_{\Theta_j} \mathcal{M}(\Theta_j, p) f(\Theta_j)$$

Assuming that the model is correct and p is the correct equilibrium parameter, we would expect to observe the frequencies of matches across agents in the population as given by $\tilde{\mathcal{M}}(p)$.[19] In general, we suggest using $\tilde{\mathcal{M}}(p)$ to test implications of the model although we reduce $\tilde{\mathcal{M}}(p)$ to a simpler form below.

5.4. **Reductions.** We will restrict ourselves to the case of $N = 4$ agents selected from the population playing the game at once. In principle, given observations regarding physical attractiveness of subjects and partners, as well as information regarding the success rates of such matches we may estimate p through $\tilde{\mathcal{M}}(p)$. This of course requires distributional

[18]We omit the notation for the conversion which is very unwieldy.

[19]We add that $\tilde{\mathcal{M}}(p)$ has no "reasonable" closed form, and accordingly the function is handled by computer simulation in this paper.

assumptions on the underlying data, and for estimation involving $\tilde{\mathcal{M}}(p)$ in entirety requires a dataset much larger than our own[20]. However, as we are actually only estimating one parameter, we may reduce $\tilde{\mathcal{M}}(p)$ to a lower dimensional form for estimation. The reduction we choose maps $\tilde{\mathcal{M}}(p)$ to the expected number of matches we should observe over the population, which we add has a nice interpretation as an efficiency measure (Aggregate Matches) which is discussed in the appendix. The reduced measure is $\mathrm{AM}(p)$:

$$\mathrm{AM}(p) \equiv \sum_{\Theta_j} [\frac{1}{4}\mathbf{1}^\top \mathcal{M}(\Theta_j, p)\mathbf{1}] f(\Theta_j) = \frac{1}{4}\mathbf{1}^\top \tilde{\mathcal{M}}(p)\mathbf{1}$$

The $\frac{1}{4}$ comes from the fact that in each game, 4 agents are playing. Division by 4 gives the expected chance of a match resulting from drawing a player from each game instance at random.

Define μ_i to be the Bernoulli parameter indicating the true rate of successful matches for type i. We are interested in the random variable, say Z, which is equal to 1 if a randomly drawn member of the empirical population has been matched and is 0 otherwise. Z is then distributed Bernoulli

$$Z \sim \begin{cases} 1, & \sum_i \mu_i g(\theta_i) \\ 0, & \sum_i (1-\mu_i) g(\theta_i) \end{cases}$$

Accordingly, let

$$\mu \equiv E[Z] = \sum_i \mu_i g(\theta_i)$$

Assuming the model is correct and that p_0 is the correct equilibrium parameter, we then have that the aggregate expected rate of matches over the population for any particular μ_0 is

$$\text{Expected Matches: } \mu_0 = \mathrm{AM}(p_0)$$

This leads us to the natural hypothesis test for any value of p_0, which we will use for estimation and confidence intervals:

$$H_0: \ p = p_0, \quad H_1: \ p \neq p_0$$

In fact, our approach will be indirect in that we may perform the much easier hypothesis test for any μ_0:

$$H_0': \ \mu = \mu_0, \quad H_1': \ \mu \neq \mu_0$$

and convert the results to the corresponding p statistics using $\mathrm{AM}^{-1}(p)$.

[20] $\tilde{\mathcal{M}}(p)$ is a 4×4 symmetric matrix and so has 10 elements to be estimated.

6. RESULTS

In this section we report our estimates, with emphasis on their interpretation in answering our two main questions.[21]

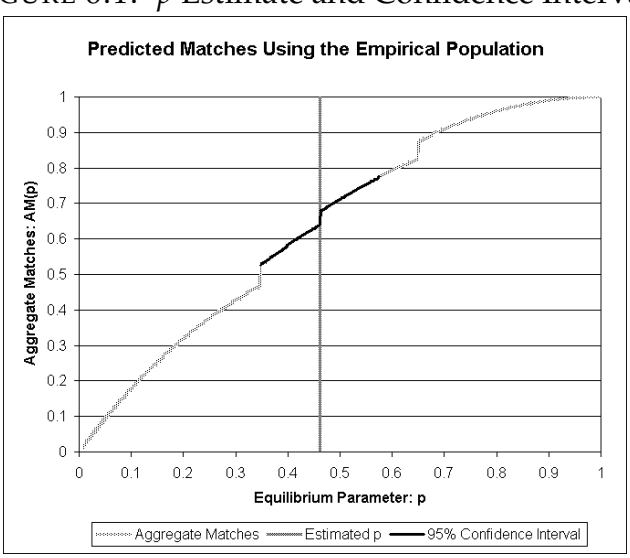
6.1. **Estimation of p.** Using the process outlined in the Methodology Section we label our estimates of μ and p as $\hat{\mu}$ and $\hat{p}$, respectively. We estimate $\hat{\mu}$ using an exact binomial test in the standard fashion. This implicitly involves estimating $\mathrm{AM}(p)$ using the population data so that $\mathrm{AM}(p)$ is in fact a (nonparametric) random variable. However, as we take the subject population of looks as the true population, we simply write $\mathrm{AM}(p)$ with the fact that it is a function of the data understood implicitly. Inverting $\mathrm{AM}(p)$ to the appropriate confidence set then yields the estimates shown in Table 5.

TABLE 5. Estimates of $\hat{\mu}$ and $\hat{p}$.

Parameter	Point Estimate	95% Confidence Interval
$\hat{\mu} = \mathrm{AM}(\hat{p})$	.648	[.506, .773]
$\hat{p}$	.462	[.348, .575]

6.2. **Interpretation.** By definition, $\hat{\mu}$ is the aggregate rate of matching for a randomly selected member of the population. More importantly, we want to understand what $\hat{p}$ says about behavior in the sex market. Our model implies that for small p, the sex market is very close to a "chasing equilibrium" and for large p, the sex market is highly assortative. We have estimated a rather midrange p, and even the ends of the confidence interval are fairly far from the extreme cases in the model as illustrated in Figure 6.1.

FIGURE 6.1. $\hat{p}$ Estimate and Confidence Interval.



[21]We note that due to a recently caught error in data entry, the reported results are slightly off (all conclusions are essentially the same). Since we are considering a change to our estimation procedure this has not been updated.

Recall that $\hat{p}$ dictates which equilibria are chosen in the sex market. Since $\hat{p}$ is an estimate, we are also interested in what the model says over the range of the confidence interval surrounding $\hat{p}$. For that purpose, we label the lower and upper confidence bounds for $\hat{p}$ as $\underline{p}$ and $\overline{p}$ respectively and examine what occurs at each value of $p \in \{\underline{p}, \hat{p}, \overline{p}\}$. The complete picture of how individuals behave in the sex market is given by who they contact given that they interact with a draw $\Theta = (t_1, t_2, t_3, t_4)$ of types from the population. We summarize each case in Table 6.

TABLE 6. Equilibria Across the Population at $\hat{p}$.

Equilibrium I	Equilibrium II	Equilibrium III	Equilibrium IV	
Θ Draws:	Θ Draws:	Θ Draws:	Θ Draws:	Θ Draws:
$(1,1,1,2)^2$	$(1,1,2,2)^2$	$(1,2,2,2)^2$	$(2,2,2,3)$	$(2,3,4,4)^1$
$(1,1,1,3)$	$(1,1,3,3)$	$(1,2,2,3)^1$	$(2,2,2,4)^1$	$(2,4,4,4)^1$
$(1,1,1,4)$	$(1,1,3,4)$	$(1,2,2,4)$	$(2,2,3,3)$	$(3,3,3,4)$
$(1,1,2,3)^2$	$(1,1,4,4)$	$(1,3,3,3)$	$(2,2,3,4)^1$	$(1,1,1,1)$
$(1,1,2,4)$	$(1,2,3,3)^2$	$(1,3,3,4)$	$(2,2,4,4)^1$	$(2,2,2,2)$
	$(1,2,3,4)^2$	$(1,3,4,4)$	$(2,3,3,3)$	$(3,3,3,3)$
	$(1,2,4,4)$	$(1,4,4,4)$	$(2,3,3,4)$	$(4,4,4,4)$

[1] indicates that for p equal to the lower confidence bound $\underline{p}$, the equilibrium is closer to I.
[2] indicates that for p equal to the upper confidence bound $\overline{p}$, the equilibrium is closer to IV.

For our estimated value of $\hat{p}$, we see that whenever multiple "low type" agents, or rather θ_1 type agents play the game, it results in equilibrium I or II indicating something like a "chasing equilibrium" where low types take their chances attempting to match with the highest types. As p increases, the equilibrium in any game tends to move towards IV, $\underline{p}$ will generally result in equilibria closer to I than under $\hat{p}$. Conversely, $\overline{p}$ results in equilibria closer to IV than $\hat{p}$. The specific movements are captured with superscripts in Table 6. In general, we see that there is a range of behavior across equilibria since $\hat{p}$ is far from the extreme values of 0 and 1. Therefore, we cannot conclude that the market is perfectly assortative, nor can we conclude that agents always chase the highest types. A more enlightening approach to analyze what occurs in the sex market is to weight the equilibria by their probability of occurrence. This results in Table 7.

TABLE 7. Distribution of Equilibria Around $\hat{p}$.

	Equilibrium I	Equilibrium II	Equilibrium III	Equilibrium IV
$\Pr(\text{Eqm}\|\underline{p})$:	.031	.05	.109	.81
$\Pr(\text{Eqm}\|\hat{p})$:	.013	.044	.105	.834
$\Pr(\text{Eqm}\|\overline{p})$:	.004	.010	.046	.94

We see from Table 7 that at $\hat{p}$, the chasing equilibrium I occurs rarely, and that approximately 83.4% of the time the game is in equilibrium IV, and therein perfectly assortative. It would be fair to say that at $\hat{p}$, the sex market is 83% assortative, 1% chasing and in intermediate cases the rest of the time. Similar interpretations can be made for the confidence bounds $\underline{p}$ and $\overline{p}$, and we see that at the 95% confidence level, our estimate indicates

that the sex market may be highly assortative indeed, where in 94% of all instances of the game, agents pair off perfectly in equilibrium IV.

7. Discussion

Our predictions regarding behavior are stochastic in nature and reflect dynamic behavior of multiple sequential partners over time. This differs in approach from what would likely be considered the most seminal pieces regarding matching behavior, namely Gale and Shapley's stable matching problems and Becker's Theory of Marriage(Gale and Shapley, 1962; Becker, 1973). In this section we highlight some of the strengths of our new approach.

7.1. **Resolving an Apparent Conflict.** The model we have chosen for analysis provides qualitative explanations of outcomes which fit with qualitative observations present in the literature. For example, Sergios and Cody (1986, pg. 72) point out the following apparent contradiction: On the one hand, Lewin's "level of aspiration" hypothesis predicts that competition in the sex market implies that "one's realistic social choices should be less socially desirable than one's fantasy choices."(Sergios and Cody, 1986) We will call this claim one. On the other hand, Sergios and Cody also point out that Berscheid (1973) found that heterosexual individuals prefer social contact with those more attractive than themselves, and in their own experimental work that physical attractiveness was the driving force behind associational choice(Bersheid and Walster, 1973). Call this claim two. We provide a model of the sex market in which competition actually results in individuals pursuing those more attractive than themselves, although that comes at a cost of remaining single. Therefore either statement one or two will be true, depending on the benefits and costs involved.

The empirical results answer the fundamental question: "When do the benefits and costs support statement one or statement two?" Statement one corresponds to a "like matches with like" situation as in the "sorting" equilibrium IV of the model. Similarly, statement two corresponds to all agents chasing the most attractive individuals as in equilibrium I. Table 7 tells us how often statement one or two holds. We conclude that statement one is the more generally true statement in regard to the mid-1990s MSM population in Chicago, or rather that the sex market is '83.4% assortative.' Even allowing for the lower bound of our estimate, $\underline{p}$, we would certainly be more inclined to agree that statement one is the best description of the Chicago sex market. However, for other populations statement two might hold more generally. One such likely population is MSM who self-select themselves into the Internet sex market as described in the next section. When does statement two hold within the population under consideration? Table 6 provides a heuristic: namely, when the sex market consists of a few very attractive individuals and many less attractive ones[22].

We add that this conflict cannot be resolved with a stable marriage problem approach, which has far too many equilibria for estimation. Nor could it be resolved using Becker's approach in the Theory of Marriage, which is a model of long term selection and matching, certainly not salient features of the population under consideration. The approach

[22]This also hints at the presence of "differentiated sex markets," e.g. the highly specialized clubs one finds in Boystown, Chicago.

here yields a way to untangle some of the qualitative statements found in the literature regarding sexual behavior and outcomes.

8. Future Directions and Conclusion

This paper is in many ways only a first step towards explaining dating and sexual behavior in a formal framework using empirical data. We provide what we think are two promising directions in order to take the next step. Finally, we add closing remarks.

8.1. **Further Empirical Work: The Generalized Model.** In order to achieve better estimation, we may generalize the above model to allow p to vary across types, with values, say $\{p_i\}$. Each p_i can be simultaneously estimated using techniques similar to those above. This presents little formal difficulty, but requires more sophisticated simulation techniques and likely a larger dataset. One promising avenue is to choose two values of p, say p_H and p_L for the highly attractive and less attractive members of the population. Then one can estimate the model on a heterosexual subpopulation that is likely to be involved in the sex market rather than the "marriage market." For instance, we suggest the single undergraduate student population or users of a computer dating site.

8.2. **Further Applied Work: Sex Markets and Epidemiology.**

"Computerized dating can save a lot of guesswork - but so can a bikini."

- Ed Parrish

One reason we are interested in sex markets is that when different sex markets are available, MSM as self-interested agents should choose to participate in markets which are to their greatest advantage. Therefore, different markets should self-select individuals with different behavioral patterns. Knowing more about a particular subpopulation, in particular its composition and risk factors for disease, allows for more tailored and effective public health policy. Two such "differentiated" sex markets are the Internet and offline markets. Papers using survey data from various sources[23] exhibit differences in the composition and risk factors involved between these two markets(Benotsch, Kalichman, and Cage, 2002; Kim, Kent, McFarland, and Klausner, 2001; Elford, Bolding, and Sherr, 2001). They generally conclude that in the Internet market relative to the offline market, participants:

(1) Are younger, but of a similar racial and educational background.
(2) Have more partners, who are more often casual.
(3) Engage in more high risk behavior, especially with regard to HIV.
(4) Are more likely to receive material benefits in return for sex.

Interestingly, each paper estimates the percentage of participants in the Internet market to be around one third of the total MSM population. We expect that participation in Internet sex markets has increased since the time the data for the above work was collected.

Using the model and estimation techniques outlined, given an appropriate dataset one should be able to describe different equilibrium behavior in the two "sex marketplaces" as well as which individuals are attracted to each. Given the results, one could design targeted interventions to those at high risk in the Internet sex market.

[23]Respectively, a San Francisco sexual health clinic, an Atlanta pride parade, and London gyms.

8.3. **Conclusion.** The techniques presented in this paper provide a method to evaluate qualitative claims regarding short term sexual behavior, as well as to examine the structure of the sexual marketplace. For our specific population, we have presented an estimated decomposition of sexual behavior and outcomes, providing a partial answer to the two basic questions asked in the introduction. We find that in the sexual marketplace defined by MSM in mid-1990's Chicago that choice of partners is highly, though not perfectly, assortative. We have also predicted conditions under which non-assortative outcomes occur, provided a qualitative explanation of why they occur, and estimated the frequency of such outcomes.

In the broader picture, we aim to help answer the classic economic question framed in this context: "How do we fulfill desire with limited resources?" We hope that this paper is a useful step in the direction of exposing the role of individual rationality in what is often left as a matter of unthinking animal drive.

References

BECKER, G. S. (1973): "A Theory of Marriage: Part I," *The Journal of Political Economy*, 81(4), 813–846.

BENOTSCH, E. G., S. KALICHMAN, AND M. CAGE (2002): "Men who have met sex partners via the Internet: Prevalence, predictors, and implications for HIV prevention," *Archives of Sexual Behavior*, 31(2), 177.

BERSHEID, E., AND A. WALSTER (1973): "The Body Image Report," *Psychology Today*, 7(July 1973), 119–131.

CAPLIN, A., AND K. ELIAZ (2003): "AIDS policy and psychology: A mechanism-design approach," *The Rand journal of economics*, 34(4), 631.

ELFORD, J., G. BOLDING, AND L. SHERR (2001): "Seeking sex on the Internet and sexual risk behaviour among gay men using London gyms.," *AIDS*, 15, 1409–1415.

GALE, D., AND L. S. SHAPLEY (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69(1), 9–15.

IMMORLICA, N., AND M. MAHDIAN (2005): "Marriage, honesty, and stability," in *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pp. 53–62. Society for Industrial and Applied Mathematics.

KIM, A., C. KENT, W. MCFARLAND, AND J. KLAUSNER (2001): "Cruising on the Internet Highway.[Miscellaneous]," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 28(1), 89–93.

KREMER, M. (1996): "Integrating Behavioral Choice into Epidemiological Models of AIDS," *The Quarterly Journal of Economics*, 111(2), 549–573.

LAUMANN, E. (2004): *The Sexual Organization of the City*. University of Chicago Press, Chicago.

LAUMANN, E. O. (1994): *The social organization of sexuality : sexual practices in the United States*. University of Chicago Press2000, Chicago.

LAUMANN, E. O., AND R. T. MICHAEL (2001): *Sex, love, and health in America : private choices and public policies*. University of Chicago Press, Chicago.

MIALON, H. (2003): "The Economics of Ecstasy. Working Paper.," .

PHILIPSON, T., R. POSNER, AND J. WRIGHT (1994): "Why AIDS prevention programs don't work," *Issues in Science and Technology*, 10(3), 33.

SERGIOS, P. A., AND J. CODY (1986): "Importance fo Physical Attractiveness and Social Assertiveness Skills in Male Homosexual Dating Behavior and Partner Selection," *Journal of Homosexuality*, 12(2)(Winter 1985/1986), 71–84.

SHIMER, R., AND L. SMITH (2000): "Assortative Matching and Search," *Econometrica*, 68(2), 343–369.

WARNER, J. (2006): "Mr. Right ... Or Mr. Right Now?," .

YEE, N. (2002): "Beyond tops and bottoms: Correlations between sex-role preference and physical preferences for partners among gay men. Unpublished manuscript.," .

APPENDIX A. THE TRAGEDY OF THE SEXUAL COMMONS

Casual observation supports the following story:

Picture a bar filled with agents of the same type who are happily contacting each other and going home with one another. In walks a handsome, buff college student who everyone takes notice of. The original agents abandon their earlier plans of contacting each other and overwhelmingly decide to chase the college student, the end result being that many of the original agents go home alone and frustrated.

In other words, the high premium attached to some individuals creates a "tragedy of the sexual commons" where not only does individual behavior lead to an inefficient group outcome, but better options may in fact decrease expected utility for some agents.

Example. Better options can hurt low types.

Consider a market populated by N agents of which $\frac{N}{2}$ are of type θ_L and $\frac{N}{2}$ are of type $\theta_H > \theta_L$. We may construct similar examples for all $p \in (0,1)$ so we arbitrarily choose $p = \frac{1}{2}$ and since decisions are invariant to scaling types by a constant we normalize θ_L to 1. It is not hard to show that as long as $\theta_H < 4$ low types all contact one another in any equilibrium, and that for $\theta_H > 4$ every low type contacts a high type in any equilibrium. The expected utility of a low type in equilibrium, given θ_H is:

$$u_L(\theta_H) = \begin{cases} \frac{3}{4}, & \theta_H < 4 \\ \frac{\theta_H}{8}, & \theta_H > 4 \end{cases}$$

So for low types, utility is constant until $\theta_H > 4$ at which point a "frenzy of chasing" occurs in which no low type will ever go home with another low type. Low types will actually do better when $\theta_H \in (1,4)$ than when $\theta_H \in (4,6)$. In the language above, as θ_H crosses 4, the game transitions (violently) from a "sorting" equilibrium to a "chasing" equilibrium.

We add that this chasing comes to the benefit of high types, as illustrated by their expected utility jump when θ_H crosses 4 and high types match with low types as a "last resort":

$$u_H(\theta_H) = \begin{cases} \frac{3}{4}\theta_H, & \theta_H < 4 \\ \frac{3}{4}\theta_H + \frac{1}{8}, & \theta_H > 4 \end{cases}$$

Finally, this example serves to show that equilibria (even RE equilibria) are not always efficient from the perspective of a social planner who places equal weights on all agent's utility. Specifically, fix $\epsilon > 0$ and suppose that $\theta_H = 4 + \epsilon$. From above, in equilibrium we have aggregate utility (AU) of

$$AU = 2 \cdot \frac{\theta_H}{8} + 2 \cdot [\frac{3}{4}\theta_H + \frac{1}{8}] = \frac{7}{4}\theta_H + \frac{1}{4} = 7\frac{1}{4} + \frac{7}{4}\epsilon$$

However, a social planner could dictate that agents of like type contact one another in which case we have

$$AU' = 2 \cdot \frac{3}{4} + 2 \cdot \frac{3}{4}\theta_H = 7\frac{1}{2} + \frac{6}{4}\epsilon$$

which narrowly beats out AU for sufficiently small ϵ.

Appendix B. Proofs

B.1. **A Lemma Regarding Optimal Behavior.** Explicit computation of $U(A_i, \mu_i)$ is quite tedious and for the empirical results we resort to computational methods. However, we can say something about the nature of an agent's decision rule. Fix an agent i and beliefs μ_i and referring to the stages of evaluating successful contacts in Algorithm 1, let S_{ij} denote the event that agent i is matched before the choice A_j is evaluated. In particular, when $i \geq j$ since S_{ij} only involves the choices of agents with higher types than agent j, S_{ij} is independent of A_j for $j \leq i$. Rewriting $U(A_i, \mu_i)$ as

$$U(A_i, \mu_i) \quad = \quad \Pr(S_{ii})\mathrm{E}[U(A_i, \mu_i)|S_{ii}] + (1 - \Pr(S_{ii}))\mathrm{E}[U(A_i, \mu_i)|S_{ii}^c]$$

we see that the value of the first term on the RHS is independent of A_j for $j \leq i$. Since $\Pr(S_{ii}^c)$ is also independent of A_j for $j \leq i$, we see that

$$\arg\max_{j \neq i} U(j, \mu_i) \quad = \quad \arg\max_{j \neq i} \mathrm{E}[U(j, \mu_i)|S_{ii}^c]$$

$\mathrm{E}[U(j, \mu_i)|S_{ii}^c]$ can be stated in words as "agent i's expected utility from contacting j conditional on the fact that a higher type has not matched with him." Calculating this utility is simple since agent i's utility from contacting j is zero unless i) his contact is successful and ii) agent j is unmatched which are events with probability p and $1 - \Pr(S_{ji})$. When these two conditions hold agent i gets a payoff of $u(\theta_j)$ so

$$\mathrm{E}[U(j, \mu_i)|S_i^c] \quad = \quad p(1 - \Pr(S_{ji}))u(\theta_j)$$

and we conclude that

$$\arg\max_{j \neq i} U(j, \mu_i) \quad = \quad \arg\max_{j \neq i} (1 - \Pr(S_{ji}))u(\theta_j)$$

For future reference, we summarize this result in Proposition .

Proposition 4. *Fix beliefs $\{\mu_k\}$ and let S_{ij} denote the event that agent i is matched before the choice A_j is evaluated. Then*

$$\arg\max_{j \neq i} U(j, \mu_i) \quad = \quad \arg\max_{j \neq i} (1 - \Pr(S_{ji}))u(\theta_j)$$

B.2. **Personal Standards.**

Proposition 5. *(Personal Standards) Let A_i^* be agent i's optimal choice in equilibrium. Then*

$$A_i^* \geq i - 1$$

In other words, an agent will never contact an agent who is more than one "level of attractiveness" below his own.

Proof. Considering Proposition 4, the result would follow if we could show that $\Pr(S_{ji}) = 0$ whenever $j < i$. This is clearly true for $i = N$ so that $A_N^* \geq N - 1$. Now if $A_k^* \geq k - 1$ $\forall k \geq m$, it follows that $\Pr(S_{j,k-1}) = 0$ $\forall j < k - 1$ so that $A_k^* \geq k - 1$ $\forall k \geq m$ implies $A_k^* \geq k - 1$ $\forall k \geq m - 1$. By induction we conclude the result. □

B.3. **Reciprocity.**

Proposition 6. *(Reciprocity) In equilibrium if a more attractive agent contacts a less attractive agent, the less attractive agent reciprocates with a contact. Formally, if $A_i = j < i$ in equilibrium then $A_j = i$.*

Proof. Suppose $A_i^* = j < i$ for some i and by the Personal Standards result we know that $A_i^* = i - 1$ and $\Pr(S_{i-1,i}) = 0$. From Proposition 4 we know that

$$A_i^* \quad \in \quad \arg\max_{j \neq i} (1 - \Pr(S_{ji}))u(\theta_j)$$

so we conclude that $(1 - \Pr(S_{i-1,i}))u(\theta_{i-1}) \geq (1 - \Pr(S_{ji}))u(\theta_j) \quad \forall j \neq i$. Combining this with $\Pr(S_{i-1,i}) = 0$ and $u(\theta_i) \geq u(\theta_{i-1})$ we have

$$u(\theta_i) \quad \geq \quad (1 - \Pr(S_{ji}))u(\theta_j) \quad \forall j \neq i$$

Since $A_i^* = i - 1$, we know that $\Pr(S_{ji}) = \Pr(S_{j,i-1}) \quad \forall j \neq i - 1$ so

$$u(\theta_i) \quad \geq \quad (1 - \Pr(S_{j,i-1}))u(\theta_j) \quad \forall j \neq i, i - 1$$

and for the result we need to show that

$$(1 - \Pr(S_{i,i-1}))u(\theta_i) \quad \geq \quad (1 - \Pr(S_{j,i-1}))u(\theta_j) \quad \forall j \neq i - 1$$

which would follow from above if we could show that $\Pr(S_{i,i-1}) = 0$. Let $k_0 \equiv \max\{k : A_k^* = k - 1\}$. By assumption such a k_0 exists and clearly $\Pr(S_{k_0,k_0-1}) = 0$ since by Personal Standards, $A_j^* > j - 1 \quad \forall j > k_0$. By above, we conclude that $A_{k_0-1}^* = k_0$. Now proceeding successively with k_t defined by $k_t \equiv \max\{k > k_{t-1} : \quad A_k^* = k - 1\}$ we conclude the result.

□

B.4. **Chasing and Sorting.**

Proposition 7. *Define the "2-person clique" function ρ by*

$$\rho(i) = \begin{cases} i + 1, & i \text{ odd} \\ i - 1, & i \text{ even} \end{cases}$$

Assume N is even. Then $\exists \epsilon_c, \epsilon_s > 0$ such that in any equilibrium

(1) *All equilibria are "chasing equilibria" if $p \in (0, \epsilon_c)$, formally $A_i = N \quad \forall i < N$.*
(2) *All equilibria are "sorting equilibria" if $p \in (1 - \epsilon_s, 1)$, formally $A_i = \rho(i)$.*

Proof. (Sketch) **Claim 1:** From Proposition 4 we know that for each equilibrium choice A_i^* we have

$$A_i^* \quad \in \quad \arg\max_{j \neq i} (1 - \Pr(S_{ji}))u(\theta_j)$$

and we know that $\Pr(S_{ji}) \searrow 0$ as $p \searrow 0$. Consequently as $p \searrow 0$, $\max_{j \neq i}(1 - \Pr(S_{ji}))u(\theta_j) \longrightarrow \max_{j \neq i} u(\theta_j)$ and by assumption all θ_i are different and u is strictly increasing so for sufficiently small p, $A_i^* = N \; \forall i < N$ which shows Claim 1.

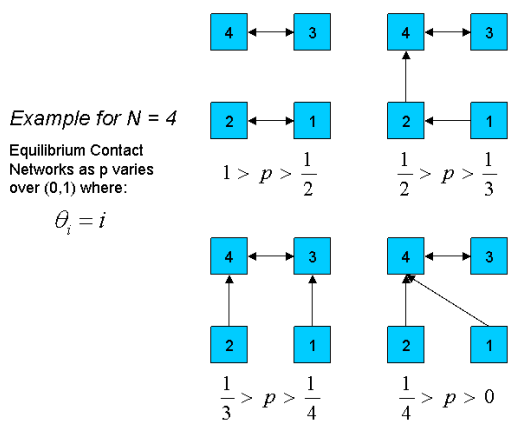
Claim 2: We know that the conclusion holds for agent N since $A_N^* = N - 1$ and for agent $N - 1$ by an application of Reciprocity. Fix $i = N - 2$ and by Personal Standards we also know that

$$A_i^* \quad \in \quad \arg \max_{j \neq i, j \geq i-1} (1 - \Pr(S_{ji}))u(\theta_j)$$

and if $j > i$ then $\Pr(S_{ji}) \nearrow 1$ as $p \nearrow 1$ so as $p \nearrow 1$ we have that $\max_{j \neq i, j \geq i-1}(1 - \Pr(S_{ji}))u(\theta_j) \longrightarrow u(\theta_{i-1})$ since $(1 - \Pr(S_{ji}))u(\theta_j) \longrightarrow 0 \quad \forall j > i$. We conclude that for a sufficiently large $p > p_i$ we have $A_i^* = i - 1$ which implies $A_{i-1}^* = i$ by Reciprocity. Setting $i = N - 2k$ for each $k > 1$ we obtain a new p_{N-2k} s.t. $\forall p \in (\max_{m \leq k}\{p_{N-2m}\}, 1)$ we have $A_{N-2k}^* = N - 2k - 1$ and $A_{N-2k-1}^* = N - 2k$. We conclude that $1 - \epsilon_s \equiv \max_m\{p_{N-2m}\}$ suffices for the result. □

Appendix C. Supplemental Appendix: Specific Model Example

Figure C.1. Equilibria Across p in a four agent game.



In this Section we illustrate the all different types of equilibria which occur as p varies as depicted in Figure C.1. Types are assumed to be $\theta_i = i$, utility is given by $u(x) = x$ and A_i^* denotes agent i's equilibrium choice. Throughout we will ignore "ties" which allow for multiple equilibria at $p \in \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\}$, as in in Figure C.1. From the Personal Standards proposition, agent 4 has $A_4^* = 3$ for all p. From the Reciprocity proposition, this implies agent 3 has $A_3^* = 4$ for all p. Now the only choices left to determine are A_2^* and A_1^*.

A_2^*: Now the probabilities of agent 4 and 3 being unmatched when agent 2's contact is evaluated are the same, namely $(1 - p)^2$. Therefore contacting agent 4 weakly dominates contacting agent 3. From Proposition 4 we know that

$$\arg \max_{j \neq 2} U(j, \mu_2) \quad = \quad \arg \max_{j \neq 2} (1 - \Pr(S_{j2}))u(\theta_j)$$

Now $A_2^* = 4$ iff

$$(1-p)^2\theta_4 = (1-\Pr(S_{42}))u(\theta_4) \quad > \quad (1-\Pr(S_{12}))u(\theta_1) = \theta_1$$

iff $p < \frac{1}{2}$.

A_1^*: Since agent 1's optimal decision depends on Agent 2's action, we break this into two cases.

Case 1: $p > \frac{1}{2}$. Now $A_2^* = 1$ so by Reciprocity, we have $A_1^* = 2$.

Case 2: $p < \frac{1}{2}$. Now $A_2^* = 4$. Appealing to Proposition 4 we have $A_1^* = 4$ iff

$$\begin{aligned}(1-p)^3\theta_4 &= (1-\Pr(S_{41}))u(\theta_4) \\ &> \max\{(1-\Pr(S_{31}))u(\theta_3), (1-\Pr(S_{21}))u(\theta_2)\} \\ &= \max\{(1-p)^2\theta_3, (1-p)\theta_2\}\end{aligned}$$

which holds iff $p < \frac{1}{4}$. Similarly, $A_1^* = 3$ iff $p \in (\frac{1}{4}, \frac{1}{3})$ and $A_1^* = 2$ iff $p > \frac{1}{3}$.

Put together these calculations generate the multiple equilibria displayed in Figure C.1.

Appendix D. Data Considerations

Here we discuss additional issues relevant to the definition of the appropriate sample not found above.

D.1. **Sexually Transmitted Infections and Relationships.** Several studies point out that the sexual behavior of those with sexually transmitted infections (STI) differs from that of the general population, likely due to risky or promiscuous behavior leading to higher rates of infection. To ensure that we are not grouping two fundamentally different populations together, we examine the rates of relevant STI found in the population. We summarize the percentages of STI in the population in Table 8 for reference and discussion.

Table 8. STI Rates.

Curable STI Rates		Incurable STI Rates	
Chlamydia	5.5%	Hepatitis B	18%
Gonorrhea	29%	Herpes	7.3%
Syphilis	1.8%	HIV	3.6%

Rates refer to whether an individual has ever been diagnosed with a particular STI. As rates are self-reported they likely underestimate true rates of incidence, although with the exception of gonorrhea, no individual was diagnosed twice for any STI. Since curable STIs can be cheaply and effectively treated once diagnosed, we do not believe the incidence of curable STIs are good grounds to separate individuals into different populations.

Hepatitis B is fairly prevalent in the population with a sufficient number of subjects to evaluate whether the knowingly infected (HB) and general population (GP) behave differently. A test for this is using part_i to see if the number of sexual partners over the course of the previous year is different across the two populations. Using a two-tailed t-test, we do not reject the hypothesis that the HB and GP populations are the same in terms of sexual behavior at the 1% level.

Herpes and HIV. As we have very few observations of individual with reported infection for these STIs, we remain agnostic about possible differences in behavior and keep these

TABLE 9. Incidence of Hepatitis B.

HB partner mean	GP partner mean
13.4	14.1
(4.98)	(4.86)

subjects in the sample. The most important of the two is HIV due to the fact that it is currently both incurable and fatal, which one might think would lead to reduced sexual behavior or monogamy uncharacteristic of the rest of the population. In fact, neither is the case, with both HIV positive individuals reporting 15 partners over the course of the past year. This indicates both are active participants of the sex market, and further supports leaving them in the sample.

D.2. **Concurrent Partners.** Several subjects were currently involved or had been involved in two or more sexual relationships concurrently. In fact, 35% of the sample had sex with someone besides the current partner during the course of their most recent relationship. In addition, 11% were involved in a sexual relationship with someone besides their most recent partner when they first slept with their most recent partner. We acknowledge these facts here as in this paper we do not examine polyamorous characteristics or model strategic behavior which takes into account polyamorous behavior. However, polyamory certainly has bearing on the questions we ask in this paper, and would be a productive avenue for further research regarding epidemiological implications.

D.3. **Data Minutiae.** We are interested in MSM who exclusively and actively have sexual relationships with men. Accordingly we dropped all female subjects and those answering anything but "Only Men" or "Mostly Men" to a question regarding gender preference. Two remaining individuals were dropped who had exclusively female partners. Another individual who answered "Don't Know" to his own attractiveness rating was dropped, as this is a primary variable of interest. Finally, two individuals who had fewer than two partners in the last five years were dropped.

Judgement Calls. As this is survey data, we dropped three other observations based on population characteristics before analysis. We acknowledge them here specifically due to the subjective nature of removal. Two individuals who answered that homosexual sex was "always wrong" to a sex attitude question were dropped both for the content of the question and because they answered 1 on the attitude scale from 1 to 4 with no 2's recorded in the sample. Another individual who uniquely reported himself as the most unattractive member of the population claimed about 10% of the most attractive sexual partners in the entire sample. In fact, unless the sex market is perfectly assortative, the model predicts such behavior, but given that there is only one individual reporting this level of attractiveness we have removed him from the data.

Variable Construction. The variable for frequency of sex over the last year must be reconstructed as pointed out in the data reference, using sex frequency for those with only one partner found in P1OFTEN and otherwise using SEXFREQ. Unfortunately the underlying variables don't exactly mesh but we construct sex frequency as detailed in Table 10.

TABLE 10. Sex Frequency Construction.

P1OFTEN		SEXFREQ
"Once a day or more"	$\longrightarrow$	1+ times a day
"4 to 6 times a week"	$\longrightarrow$	3-6 times a week
"1 to 3 times a week"	$\longrightarrow$	1-2 times a week
"2 to 3 times a month"	$\longrightarrow$	2-3 times a month
"About once a month"	$\longrightarrow$	1 time a month or less
"Once or twice"	$\longrightarrow$	1-2 times a year