

## Degree Distributions in Sexual Networks: A Framework for Evaluating Evidence

DEVEN T. HAMILTON, MPH,\* MARK S. HANDCOCK, PhD,† AND MARTINA MORRIS, PhD\*‡

**Objective:** We present a likelihood based statistical framework to test the fit of power-law and alternative social process models for the degree distribution, and derive the sexually transmitted infection epidemic predictions from each model.

**Study Design:** Five surveys from the United States are analyzed. Model fit is formally compared via Akaike Information Criterion and Bayesian Information Criterion, and substantively assessed via the prediction of a generalized epidemic.

**Results:** Formal goodness-of-fit tests do not consistently identify any model as the best all around fit to the US data. Power-law models predict a generalized sexually transmitted infection epidemic in the United States, while most alternative models do not.

**Conclusions:** Power-law models do not fit the data better than alternative models, and they consistently make inaccurate epidemic predictions. Better models are needed to represent the behavioral basis of sexual networks and the structures that result, if these data are to be used for disease transmission modeling.

THE RAPID BUT UNEVEN GLOBAL spread of human immunodeficiency virus (HIV) has challenged researchers to better understand sexually transmitted infection (STI) transmission dynamics and their implications for prevention. The basic components of transmission are captured by the well-known population summary measure, “ $R_0$ ,” the reproduction rate or reproduction ratio of infection. The reproductive rate is the number of secondary cases generated from a single infective case introduced into a susceptible population.

$$R_0 = \tau c \delta$$

where  $\tau$  = the probability of transmission per contact;  $c$  = the contact rate; and  $\delta$  = the duration of infection.

If  $R_0$  is  $<1$  the initial infected case produces  $<1$  new infection on average, transmission is below the reproductive threshold and the disease is unlikely to persist.

Early work in the population dynamics of STI focused on the apparent paradox that transmission persisted despite contact rates in the general population that suggested an  $R_0$  below the epidemic threshold. The result was the development of the core group concept, the theory that a small group of people with large numbers of partners and repeated infections can maintain a reservoir of infection within a population even when the average reproduction rate for the population is below threshold.<sup>1,2</sup> Such insights directed

*From the Departments of \*Sociology and †Statistics, University of Washington, Shoreline, Washington*

both policy makers and health-care providers to focus on identifying and treating individuals with repeat infections to reduce  $R_0$  within this highly active core group, a strategy that seeks to curtail disease spread without the cost and coordination challenges of broad population-based screening and intervention. Central to these prevention efforts was the ability to use repeat infection as a reliable marker for core group membership, allowing for accurate intervention targeting. Treatment seeking behavior by those repeat infections also created a natural synergy between identification and treatment efforts.

HIV does not fit into the core group framework, in part because there is no reinfection (so no simple marker for risk and no readily identifiable core group), and in part because of the generalized epidemics of HIV in many countries (which suggests transmission is above the reproductive threshold in large sections of the population).

An epidemic is formally defined as generalized once 1% of the general population is infected,<sup>3</sup> i.e., once it has spread broadly beyond groups with traditionally defined risk factors such as injection drug use, commercial sex, other sexually transmitted infections, or large numbers of sex partners. These markers of risk have typically been used to target prevention strategies, but with HIV prevalence above 30% in some countries (not just their capital cities), the concept of a “core group” begins to lose meaning and relevance. Prevention strategies are therefore being rethought in the context of HIV. If a small core group is not responsible for persistence, then first principles dictate going back to think about the mechanism: the transmission network. The integration of network analytic theory and methods into HIV prevention research has provided a number of new insights into STI population dynamics.<sup>4</sup>

In the network context, it is also possible to directly calculate  $R_0$  under some circumstances. Under the assumption of random partnership formation given the heterogeneity in contact rates across persons<sup>5</sup>:

$$R_0 = \frac{\bar{\tau}}{\tau_c}$$

where  $\bar{\tau}$  = the average integrated probability of transmission given a partnership; and  $1/\tau_c$  = the contact network epidemic potential.

The “average integrated transmissibility” is a partnership level measure that represents the likelihood of transmission in a discordant pair, averaged across all such pairs in the population. It is determined

Supported by NIH grants R29HD034957 and R01DA012831.  
Correspondence: Deven Hamilton, MPH, Department of Sociology, University of Washington, Box 353340, Seattle, WA 98195. E-mail: dth2@u.washington.edu.  
Received for publication January 3, 2007, and accepted June 12, 2007.

by the characteristics of partnerships (e.g., duration, coital frequency), the characteristics of the pathogen (transmissibility), and the characteristics of the host (infectivity and susceptibility). We will refer to it below by the simpler term “transmissibility.” The “contact network epidemic potential” is a population level measure that summarizes the level of connectivity in the contact network. It is determined by the network structure (e.g., the density, distribution, and clustering of partnerships), and its form depends on the model for the network.<sup>6</sup> As the form of the equation suggests,  $\tau_c$  can also be interpreted as the threshold transmissibility: the transmissibility needed to sustain an epidemic in this particular network. We will refer to these below as “epidemic potential” ( $1/\tau_c$ ) and “threshold transmissibility” ( $\tau_c$ ).

The key feature of this generalized  $R_0$  is that neither transmissibility nor the contact network alone determine the likelihood of an epidemic; it is the relationship between the 2 that is important. Because of this interdependence, an extremely high level of transmissibility may not induce an epidemic if the epidemic potential of the contact network is sufficiently low. By the same token, very low levels of transmissibility may induce an epidemic if the potential is sufficiently high. An extreme case that has garnered a great deal of attention in the recent literature is the case of so called “scale-free” networks. In this case, the theory defines  $\tau_c = 0$ , so there is no epidemic threshold. Under these conditions any level of transmissibility is sufficient to induce an epidemic with very high probability. The speed and the magnitude of the epidemic will be determined by the level of transmissibility. The scale-free case is, however, a single extreme state of the degree distribution; the theory of network mediated transmission is considerably more general, and more empirically useful.

In this article, we present a statistical framework that allows us to test the fit of alternative models for the degree distribution, estimate the resulting epidemic potential of the network, and derive the epidemic predictions of the models in terms of  $R_0$ . We use these methods to analyze 5 different surveys of the US population. Using multiple surveys of the same population helps to highlight how robust the statistical estimates and epidemic predictions are to subtle changes in the survey sample, questionnaire, and data collection method.

The original expression for  $R_0$  represents the impact of sexual activity via the single homogeneous term,  $c$ . The simplest generalization of this term is to allow for heterogeneity in contact rates across persons: the population “degree distribution.”<sup>5,7,8</sup> In a sexual network, the degree,  $K$ , is the number of sexual partners of a randomly chosen member of the network and the distribution of  $K$  is referred to as the *degree distribution*. Empirical evidence shows that degree distributions for sexual partnership networks are highly right skewed.<sup>5</sup> Although the long upper tail is important for transmission, the other feature of these skewed distributions is a very dense lower tail: the modal degree is 1 partner in the last 12 months for nearly all large representative surveys (e.g., Refs. 9–16). In the 5 survey data sets that will be used in our analysis, 64% to 76% of respondents report exactly 1 partner in the last year.

The extreme right skew of these degree distributions has led researchers to explore whether these distributions can be represented by a “power-law” scaling function that is found in many other physical systems. The probability mass function  $P(K = k)$  of a degree distribution has power law behavior with scaling exponent  $\rho$  if there exist finite constants  $c_1$ ,  $c_2$ , and  $M$  such that  $0 < c_1 < P(K = k)k^\rho < c_2$  for  $k > M$ . The relevant characteristic of these models for understanding disease transmission is the scaling parameter ( $\rho$ ). If this parameter is between 2 and 3, the variance of the degree distribution is theoretically infinite and the distribution is said to be “scale free.” With infinite variance the epidemic potential,

$1/\tau_c$ , becomes infinite, so the infectivity  $\bar{\tau}$  needed for an epidemic approaches zero. That is, an epidemic is a virtual certainty for any pathogen that has a nonzero probability of transmission. Concomitantly, the critical vaccination fraction under such circumstances is unity. There is no herd immunity—all members of the population must be successfully vaccinated to eradicate the disease. Under these conditions, only interventions that successfully shut down the high-degree “hub” nodes will restore the epidemic threshold and curtail the diffusion process.<sup>17</sup>

The scaling parameters have been estimated by several researchers across a range of populations including Uganda the United States, Sweden, Britain, and Zimbabwe. The results have been somewhat inconsistent and often counterintuitive. The results suggest there should be large STI and HIV epidemics in Britain, Zimbabwe, and potentially the United States, but not in Sweden or Uganda.<sup>5,7,8,18,19</sup> The results are summarized in Table 1. The discontinuity between the model predictions and the empirical data serve as the motivation for the current investigation.

Why do these model estimates fail to predict results consistent with the evidence? It is our contention that the methods employed in these studies have been inadequate in 1 or more of the following ways. First, the amount of error in the tail of the distribution has been underestimated. Second, models for the social process have been bypassed in favor of simple curve fitting. Finally, models that focus exclusively on degree distributions ignore heterogeneous population mixing patterns, so are unlikely to provide informative predictions of epidemic potential. As a consequence, the evidence that power law models adequately fit the degree distributions of empirical sexual networks has been overstated.

The typical approach to estimating the scaling parameter  $\rho$  of a power-law model is to fit a regression line through the apparently linear region of a plot of the survival function of the degree distribution plotted against the distribution on double logarithmic axes<sup>7,18</sup> and the standard error of the estimated slope is used as an estimate of uncertainty in the model. The appeal of this approach is due to its familiarity and simplicity. There is, however, an implicit assumption here that measurement uncertainty is not correlated with degree. The data suggest that this assumption is not met, for 2 reasons. One is that there is little information in the tail of the distribution, since so few people report large numbers of partners. In the data sets we analyze here, for example, <1% of respondents report >10 partners. Another source of error in the tail is response rounding. Above 10 partners there is a tendency for respondents to report in round numbers, and above 20 partners over 80% of the reports are in large round numbers.<sup>20,21</sup> As a consequence, the precision of the model approximation becomes more dependent on less accurate data as it fits the higher degree values. And because the distribution tail has a disproportionate effect on the slope of any line fit, the scaling parameter estimates tend to be very unstable. Approaches that use data only from the upper tail of the distribution<sup>8</sup> simply exacerbate the problem. In this context, the simple regression methodology is inappropriate. It yields biased estimates of the scaling parameter, and greatly underestimates model uncertainty.<sup>22</sup> A more sophisticated form of curve fitting is used in.<sup>8</sup> They fit a curve to  $P(K = k)$  of the form  $ck^{-\rho}$  using a form of maximum likelihood. It has the disadvantage of neither using a social process model nor adjusting for the selection of the upper tail as is done here.

The curve fit approach raises another problem as well—the assessment of alternative hypotheses regarding the underlying process that gives rise to the observed degree distribution. This kind of statistical evaluation requires a theory of the mechanism of partner acquisition, represented in the form of a stochastic model that can then be fit to the data and compared to alternatives. As

TABLE 1. Reported Power-Law Scaling Parameters and Confidence Intervals From Studies of the Degree Distributions of Sexual Networks in 5 Countries

Country	Authors	Data Source	Study Population	$\rho$ Estimate
England	Schneeberger et al., 2004	National Survey of Sexual Attitudes and Lifestyles (1990 and 2000)	A household-based probability sample of adults in Britain	1990—males: 2.85*, CI (2.8–2.91); females: 3.68, CI (3.59–3.78) 2000—males: 2.48*, CI (2.43–2.53); females: 3.1, CI (3.03–3.17)
		London Gay Men's Sexual Health Survey	A convenience sample of gay men in London	Males: 1.87*, CI (1.8–1.94)
Sweden	Handcock and Jones 2004	Sex in Sweden Survey (1996)	National probability sample of adults in Sweden	Males: 3.25, CI (3.01–3.63); females: 4.23, CI (3.60–5.21)
Uganda	Handcock and Jones, 2004	Rakai Project Sexual Network Survey (1995)	Rakai District adults 16–45	Males: 5.43, CI (4.32–6.53); females: 17.04, CI (12.58–25.19)
United States	Handcock and Jones, 2004	National Health and Social Life Survey (1992)	Nationally representative sample of US adults 18–55	Males: 3.03†, CI (2.8–3.32); females: 3.84, CI (3.34–4.55)
Zimbabwe	Schneeberger et al., 2004	Population-based survey of rural Zimbabwe reported by Gregson et al. 2000	Adults in rural Zimbabwe	Males: 3.07†, CI (2.87–3.29); females: 2.51*, CI (2.12–2.97)

\*Scaling parameters imply a degree distribution with infinite variance.

†Point estimates >3, but with confidence intervals that include values <3.

others have noted, power law behavior is consistent with a “preferential attachment” mechanism.<sup>23–26</sup> In the context of sexual partner selection, this would imply that people chose partners with a probability proportional to the partner's degree: the more partners that person has, the more likely you would be to choose them. Although this may be a plausible mechanism for link formation in the World Wide Web—where the goal is to minimize steps to all other nodes—it is a questionable representation of how people choose their sexual partners. Indeed most social norms would suggest the opposite; that having many partners makes one socially undesirable. In any case, to move this discussion from speculation to empirical assessment, it is necessary to define and test alternative models of the process.

We identify several plausible alternative models and compare how well they fit the data from 5 surveys. In testing several alternative models on several different data sets, we hope to gain some leverage on identifying the range of processes that describe partnership distributions. If one model fits better than the others in a wide variety of cases, this suggests a commonality of processes. If the best-fit model varies from one data set to the next, this suggests that the partnership acquisition process is too complex to capture with a simple model based exclusively on degree. In that case, the reported fits of individual models would most likely be idiosyncratic and not particularly informative. In addition, if different models fit the different measures from the same data set, this would suggest the results are very sensitive to the measure used, and this should limit the inferences made. Finally, the sample data themselves may be inadequate and fail to accurately capture the true degree distribution in the population. Previous research however, has shown remarkable consistency in reported sexual behavior<sup>27</sup> and there is no a priori reason to believe the data are not representative.

Regardless of which model fits the data best, a good model must also predict the epidemic behavior we empirically observe. The ultimate goal of analyzing degree distributions is to develop a methodology that sheds light on disease transmission dynamics in

different populations. So in addition to comparing the fit of each model, we also examine the epidemic potentials predicted by each.<sup>5</sup> If the best fitting model for the degree distribution still produces nonsensical epidemic predictions, this again implies that a degree-based model is too simple to capture the true underlying transmission networks.

## Materials and Methods

### Data

The degree distribution measures we will examine are drawn from questions about the number of partners in the last year. It is worth a brief discussion on the rationale for that choice. The appropriate time slice to use for estimating network epidemic potential is not obvious, an issue often ignored in the literature. Because partnerships have duration and sequence, the true transmission network does not exist during a single time period, but is instead a concatenation of periods that defines the possible infection path, given the duration (and/or stages) of infection. Outside of computer simulations, we generally can not observe this network. Our goal here, however, is to capture the degree distribution this transmission network would have if we could see it. Consider the degree distributions based on lifetime partners and current partners. Both of these choices are useful in certain contexts, but neither is likely to provide a good representation of the underlying transmission network. The use of lifetime partnerships ignores the impact of partnership sequence and dramatically overestimates the density of the transmission network. Current partnerships select disproportionately on long relationships and are likely to underestimate the density of a transmission network.

Using partners over the last year as a proxy for the transmission network is not perfect, but it has some merits. The longer-term partnerships are balanced by observation of some shorter-term partnerships, and the time window is a reasonable compromise for a range of STI infection durations. Treated STIs probably have shorter infection windows than a year, but this may be balanced by

TABLE 2. An Overview of the 5 Population Based Surveys Used in This Study

Survey	Years	Age	Sex	Interview Method	Question	Data Adjustments
BRFS	1996–2000	18+	M/F	Telephone	During the past 12 mo, with how many people have you had sexual intercourse?	No adjustments. Responses top-coded at 76+. Three respondents reported 76+ partners
GSS	1988–1991, 1993, 1994, 1996, 1998, 2000	18+	M/F	Self-administered questionnaire	How many sex partners have you had in the last 12 mo?	Responses were categorical for values greater than 4 so all responses greater than 4 were coded as 5. These data were only used in the truncated analysis
NHSLS	1992	18–59	M/F	Face-to-face interview	Thinking back over the past 12 mo, how many people, including men and women, have you had sexual activity with, even if only 1 time?	An additional question was asked on the SAQ portion of the questionnaire which is also used as a variable constructed by the original researchers
NSM	1991	19–41	M	Face-to-face interview	Since January 1990, how many different women have you had vaginal intercourse with? Since January 1990, how many different partners have you had anal sex with?	Constructed from the greater of vaginal sex partners in 1990 and anal sex partners in 1990
NSW	1991	19–38	F	Face-to-face interview	With how many different men did you have vaginal intercourse since January 1990? With how many different men did you have vaginal intercourse since January 1990?	Constructed as the greater of vaginal sex partners in 1990 or anal sex partners in 1990

some level of asymptomatic and thus longer infections. Untreatable infections often have periods of peak infectivity, like recent infection with HIV,<sup>28</sup> and the flare-ups of herpes simplex virus. Finally, the yearly degree distribution allows for comparability with previous literature, and it is a time period commonly used in sexual behavior surveys. Further discussion of this issue is given in Ref. 59.

We analyze data from 5 different sexual behavior surveys with probability samples collected in the United States during the 1990s: the National Health and Social Life Survey (NHSLS), the General Social Survey (GSS), the National Survey of Men (NSM) and the National Survey of Women (NSW) and the Behavioral Risk Factor Surveillance Survey (BRFS). A brief overview of the data are provided in Table 2.

The GSS is conducted by the National Opinion Research Center (NORC) and is designed as part of an ongoing program of social indicator research to gather repeated measures on a broad range of data. The GSS uses the NORC national probability sample, which includes all noninstitutionalized English-speaking persons 18 years of age or older living in the United States. The samples are designed to give each English-speaking household in the United States an equal probability of inclusion. The GSS asked questions about the number of sexual partners in the last year in 1988–1991, 1993, 1994, 1996, 1998 and 2000. Pooling across years, there is a total of 16,159 respondents who reported the number of sexual partners in the last year.<sup>14</sup> GSS response codes for the number of sex partners in the last year are categorical and top-coded (1, 2, 3, 4, 5–10, 11–21, 21–100, 100+).

The NHSLS is a comprehensive survey of the sexual behavior of US adults 18 to 59 conducted in 1992 by NORC.<sup>29</sup> The survey employed a multistage area stratified probability sample designed

to give each household an equal probability of inclusion. The sample size is 3332. Several different measures of the number of partners in the last year are available in this survey: direct response to an interviewer, response to a self-administered questionnaire, and a constructed variable provided by the study investigators that is based on consistency checks across several measures. These 3 measures will allow us to determine whether model fits are robust to small changes in reporting within the same survey.

The NSM was conducted in 1991 and was designed to examine sexual behavior and condom use among young adult men 20 to 39 years old. The survey was based on a multistage, stratified, clustered, disproportionate-area probability sample of households within the contiguous United States.<sup>15</sup> The sample includes 3321 noninstitutionalized men. Respondents are asked to report the number of vaginal sex partners and anal sex partners in separate questions. There is no way to ascertain how many partners are represented in both categories, so we define the number of partners as the maximum of the 2 categories, which may be lower than the actual number of unique partners. A total of 586 (19%) of the men reported anal sex. Of these, 18 report no vaginal sex partners, and 35 report more anal than vaginal sex partners.

The NSW was also conducted in 1991 and was designed to examine sexual, contraceptive, and fertility behaviors among young adult women 20 to 29 years old. The sample includes 1669 respondents from 2 subsamples. The first subsample ( $n = 929$ ) consisted of follow-up cases from the 1983 National Survey of Unmarried Women, which surveyed 1314 never-married women between 20 and 29 years of age. The second subsample ( $n = 740$ ) is from a different probability sample of 20 to 27 year old women of unspecified marital status selected in 1991.<sup>16</sup> The NSW uses a



very similar instrument to the NSM, so the same adjustment strategy is used.

The BRFS is a part of the state-based Behavioral Risk Factor Surveillance System initiated in 1984 by the Center for Disease Control to collect data on risk behaviors and preventative health practices of the US population age >18.<sup>14</sup> The BRFS includes questions about sexual behavior as part of a supplement started in 1996. States make the decision whether to include the supplement in each year. We use data for the 5 years from 1996 to 2000. The number of states electing to include the supplement during this period varied from a low of 2 in 1996 to a high of 25 in 1997. The variation makes it impossible to aggregate these data into a true national probability sample. We did not want to exclude the BRFS since it has a large sample size ( $n = 72,280$ ), but the results should be interpreted with care. The variable for the number of sex partners in the last year is top coded at 76+. Only 3 respondents reported the top-coded value.

In order to make these data comparable, we restrict our focus to the age range common to all surveys, 20 to 39. In addition, each of the data sets is resampled with an adjusted weight that combines both the post stratification weight and a renormalizing factor so that all of the data sets have the same age distribution in this range as the United States in 2000. Since the variable of interest here is partners in the last year, rather than partners over the lifetime, the restriction to the 20 to 39 range is likely to understate the fraction of the overall population with 0 or 1 partner. Data from the GSS, for example, indicate that only 8% of the population in this age range had no partners in the last year compared to 30% of respondents outside this range. Similarly, 20% of respondents aged 20 to 39 reported >1 partner in the last year compared to <10% of respondents outside this age range.

Once the data are adjusted to make them more comparable we have 12 different distributions to analyze. The NHSLs has 3 different measures of partners in the last year, so contributes a total of 6 distributions (3 measures by 2 sexes). The remaining surveys have 1 measure. GSS and BRFS each provide 2 distributions, 1 for each sex, and the NSM and NSW each provide 1 distribution for males and females respectively.

### *Social Process Models for Partnership Formation*

We examine 4 general types of mechanisms that may describe the process of partnership formation: fixed rate, a search process with a stopping rule, preferential attachment, and a two-stage vetting process. In each case, there is >1 model that can be defined for that mechanism, so in all, we examine 11 different models: the Poisson, Yule, Waring, Discrete Pareto, Negative Binomial, Geometric, Discrete Pareto Exponential, Geometric Yule, Negative Binomial Yule, Anchored Negative Binomial and the Poisson-lognormal. These models are drawn largely from<sup>5</sup> but also include the Poisson-lognormal model described by.<sup>30</sup>

The Poisson is a fixed rate model and is the natural null hypothesis in this context. All persons are assumed to acquire partners at a fixed homogeneous rate ( $\lambda$ ) over time. The assumption that everyone in the population has the same propensity to form ties is both strong and unrealistic. We do not expect this model to fit the data, but it is the most appropriate basis for comparison.

The Poisson lognormal and Negative Binomial models are simple generalizations of the Poisson, retaining the assumption that people have a fixed rate of partner acquisition, but relaxing the homogeneity assumption. Individual propensities to form ties ( $\lambda_i$ ) are drawn from a lognormal distribution (for the Poisson-lognormal) or a  $\gamma$  distribution (for the Negative Binomial). The Negative Binomial is sometimes referred to as a Poisson- $\gamma$  mixture model

and is often used in the biologic sciences. Both models are used to capture overdispersion relative to the Poisson.

The Negative Binomial and Geometric can also be interpreted as search process (or waiting time) models with stopping rules. For both, the process is represented as a sequential search with parameters ( $n, r, p$ ): over  $n$  trials, partners are acquired with probability  $p$  until the search is stopped when  $r$  suitable partners are found. For the geometric model,  $r = 1$ , and the search stops after the first suitable partner is found. For the Negative Binomial model the search stops after the specified number ( $r$ ) of suitable partners are found.

The Yule<sup>5,23</sup> and the Waring<sup>24</sup> are degree distributions that result from preferential attachment models.<sup>25</sup> Under these models the degree distributions are generated as the accumulation of partners over time where the probability that a contact is made with any particular individual is a function of that individual's current degree. These preferential attachment models are often referred to as "the rich get richer" models. The Yule model also has an alternative interpretation, as an exponential mixture of geometric distributions. The Waring is a generalization of the Yule that includes an additional parameter for the probability of recruiting an individual with no partners into the network. We include the Discrete Pareto here for comparisons with previous literature. The Discrete Pareto is the power law model with  $P(K = k) \propto k^{-p}$ , often found in the physics literature. It appears to lack a plausible stochastic mechanism for partnership formation, so is perhaps best understood as an approximation to the Yule. The "power law" specifications used in the prior literature are typically defined over the positive integers. They therefore exclude zero, although the latter is a valid response, which is often the modal value in the data.

The Negative Binomial-Yule, Geometric-Yule and Discrete Pareto Exponential are vetting models that represent partnership formation as a two-stage process. First, a person generates an acquaintance list that serves as the eligible partner pool. People then choose their sexual partners from the acquaintance list. This class of model is extremely flexible in that practically any probability distribution can be specified for each stage of the process. The vetting process may be used to represent a latent clustering of potential partners due to social networking, geographic, temporal, or other factors. The mechanism can also be used to represent a selection process designed to satisfy multiple criteria. Individuals may choose partners with a fixed probability from acquaintances that independently satisfy some criteria, for example, sex, age, marital status, and sexual preference. The accumulation process continues until they meet a quota of people that satisfy the criteria. The Yule-vetting models can also be interpreted as generalizations of the Yule distribution that recognize that the formation of sexual partnerships is not costless.

All of these models can be generalized to allow for the possibility that the process may provide poor predictions for the lower degrees, but still accurately predict the tail behavior. To isolate the upper regions of the degree distribution we allow separate parameters to fit the probabilities of lower degrees (below a given cutoff value) and fit the parametric model only to those values at or above this cutoff. This is particularly useful for comparisons to the power-law models, since they do not make predictions for  $k = 0$ .

If the cutoff is above zero the Geometric, Negative Binomial, and Poisson models start their search process after the cutoff partner. As an alternative to this, the last model, the Anchored Negative Binomial fits the values below the cutoff with additional parameters but assumes the search process started with the first partner regardless of the cutoff above one. For a more in-depth description of all of these models, see Refs. 5 and 30.

*Model Estimation*

We use maximum likelihood to estimate the parameters of these models from the 5 different survey data sets. Maximum likelihood estimation (MLE) has many desirable statistical properties in this context.<sup>31</sup> The likelihood functions are computed from the probability mass functions of the models. If the observations were categorical or top-coded, then the likelihood is modified to reflect this. For example, if the degree of an individual is only known to be between 5 and 10, then the likelihood of observing this is the sum of the probabilities of observing a degree of exactly 5, a degree of exactly 6, . . . , and a degree of exactly 10. In this way, the probabilities of categorical or top-coded observations can be derived from the probability mass functions of each degree. The resulting likelihood function is referred to as a grouped likelihood as the individual's probabilities are grouped together.<sup>22</sup>

Since the power-law models do not provide predictions for  $k = 0$ , model comparisons can only be based on the fits for degrees above zero. We use 3 cutoff values in our tests: 1, 2, and 3. This produces a total of  $33 = 11 \times 3$  models per distribution. The number of parameters for a model,  $d$ , is the number of parameters when the cutoff is zero plus the cutoff value. When using a MLE the natural comparison between 2 nested models is a likelihood ratio test. For nonnested models, as most of our comparisons are, there are several alternative methods for comparing goodness of fit. We use 2: (1) the Akaike Information Criterion (AIC)<sup>32,33</sup> and (2) the Bayesian Information Criterion (BIC).<sup>34</sup> For a simple random sample of  $n$  people with data  $K_1, \dots, K_n$ , the AIC is defined as  $AIC = -2\log \text{likelihood}[K_1 = k_1, \dots, K_n = k_n] + 2d$ , and  $BIC = -2\log \text{likelihood}[K_1 = k_1, \dots, K_n = k_n] + d \log n$ . The 2 approaches are similar but the BIC has the benefit of incorporating model uncertainty and sample size into the decision. The AIC has the advantage of efficiency.<sup>33</sup> If the complexity of the true model does not increase with the size of the data set, the BIC is usually preferred, otherwise AIC is preferred. For both, a well-fitting model will be indicated by a low value. When comparing models using BIC, the general rule is that a difference of 2 to 6 is moderate evidence of superior fit, 6 to 10 is strong evidence, and 10+ is very strong.

**Results**

We start by employing the simple curve-fitting approach used in the previous literature to estimate the scaling parameter for a power-law model. This is a helpful benchmark estimate, establishing what a naïve analysis of the degree distribution from the 5

surveys would produce. For samples with degree  $>1$  the estimates of the scaling parameters are below 3 in 10 of the 12 fits, implying the distributions are scale free. In the other 2 cases, the confidence interval generated by the standard error incorporates values that indicate a scale-free distribution. When we restrict the sample to degree  $>2$ , all but 1 of the scaling parameter estimates are below three. This approach consistently produces estimates that predict a generalized HIV/AIDS epidemic in the United States, as well as generalized epidemics of all other STI.

*Comparisons Based on Goodness of Fit*

Using MLE to fit the 11 social process models to these data, no one model emerges as the best fit across the different surveys. Table 3 shows the top 3 model fits to each data set as ranked by BIC. The top-ranked model reports the associated BIC; the second and third best fits show the sequential differences in the BIC. Scanning across the different samples in Table 3, the variation in model ranking is clear. Four of the 11 proposed models rank as the best fit in at least 1 instance across the 12 distributions; 5 of these are variants of the power law models, 5 are a search/stop rule model (the Negative Binomial), and 2 are a vetting model (the Negative Binomial Yule). Only 6 of the 11 models ever show up the top 3 ranking. Of these, the majority are power law (19 of the 36); the remainder nearly evenly split between vetting (10) and search/stop rules (7). One model that never shows up in the top 3 ranking for any data set is the null—the Poisson model of a single homogeneous rate of partnership formation over time.

The difference in BIC across the top 3 ranked models for any survey is typically large enough to indicate strong or very strong evidence of superior fit at each rank. An example of the fits of the different models to the observed data are illustrated in Figure 1, which shows the predicted distributions under each model plotted over the observed data from the NSM and NSW. The fit of several different models to the same data is almost indistinguishable. This suggests that the information in the distributions can not discriminate among these models.

The variations in fit display some systematic patterns by sex. For the male degree distributions, the top-ranked models are mostly search/stop rule and vetting models, though power-law models account for half of the models in the top three ranking (9 out of 18). For the females, by contrast, the power-law models are slightly more likely to provide the best fit: 4 of the top 6, and 10 of the 18 top three.

There are also systematic differences in model fits by survey. The GSS and 2 of the 3 distributions from the NHSLs (categorical

TABLE 3. The 3 Best-Fitting Models to Each of the 6 Distributions in Order of Rank By BIC

	Rank of Fit	NHSLs Constructed	NHSLs Continuous	NHSLs Categorical	GSS	BRFS	NSM-NSW
Males	1	NB 3,328.4	DP* 3,446.5	NB 3,225.4	NB 7,425.1	NBY 40,891.2	NBY 9,024.5
	2	NBY 62.7	War 6.4	DP 60.8	DP 274.1	NB 61.6	NB 20.5
	3	War 82.4	DPE 7.1	War 65.4	War 275.4	War 85.5	War 26.5
Females	1	DP 3,264.4	NB 3,113.1	DP 3,240.9	War 6,751.8	War 35,559.9	NB 3,884.0
	2	War 0.1	War 17.9	War 4.5	DP 4.2	DP 10.2	DPE 1.0
	3	NBY 5.3	NBY 18.7	NBY 15.3	Yule 4.9	NBY 14.1	NBY 8.5

Dark grey, power law; light grey, vetting; white, search/stoprule.

1st ranked: NB, DP, NBY, War.

2/3 ranked: DPE, Yule.

NB indicates negative binomial; DP, discrete pareto; NBY, negative binomial yule; War, Waring; DPE, discrete pareto-exponential.

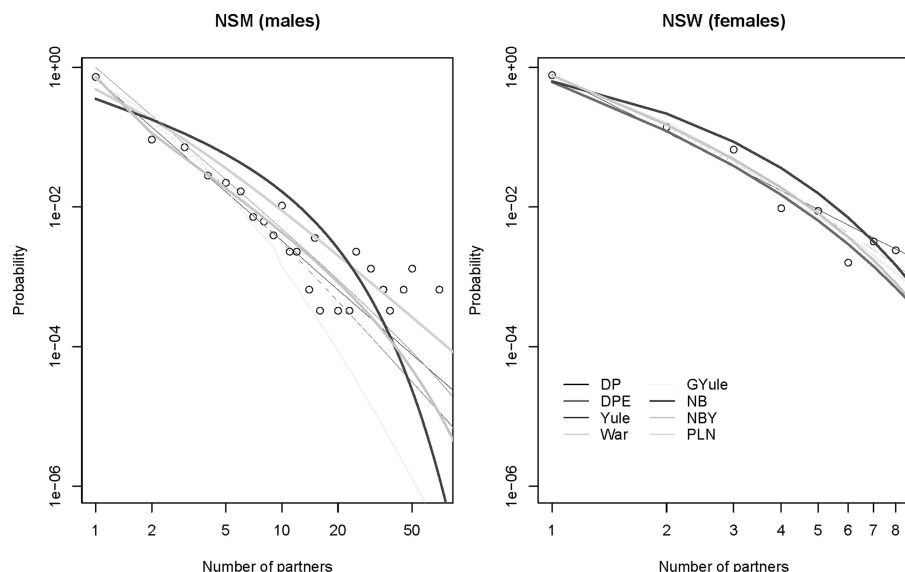


Fig. 1. The predicted probability mass functions from 8 models fit to the observed data from a survey of US men and women aged 20 to 39 years.

and constructed) produce identical model rankings for men and nearly identical rankings for women. The ranking is somewhat different for men and women but both have the Discrete Pareto and Waring power law models in the top 3. The similarity in the fits is not altogether surprising—the two NHLS distributions come from different measures in the same survey and both the GSS and NHLS use a true nationally representative sampling frame drawn by the same organization, NORC. What is surprising is that the continuous measure from the NHLS is best fit by a different set of models. This indicates that the model fits are not robust to changes in question design, even within the same survey. The BRFS and NSM have identical model rankings for men, though not for women, and again share a number of models in the top 3.

How robust are these findings to fits based on the upper tail of the survey distributions? The AIC, BIC and log-likelihood for the 3 best-fit models to all 12 distributions at the three different cutoffs are shown in Appendix 1. Here again, only a few models find their way into the top rank. Five of the 11 proposed models rank as best at least once among the 24 distributions defined by the 2 upper-tail cutoffs: Half are power-law models; most of the rest are search/stop rule models. Nine of the 11 models eventually show up in the top 3: Half are power-law models, and the other half are split between search/stop rule and vetting, with a few Poisson-lognormal (PLN) mixtures.

There is little consistency in the specific model rankings within a survey across the cutoff values, but there is some consistency in the types of models that fit across cutoffs. If a power law model fit the distribution best for the full distribution, a power law also tended to fit the distribution best at cutoffs 2 and 3. And full distributions best fit by search/stop rule models were also fit best by these at cutoffs 2 and 3. The NSM was an exception; for this survey, there was almost no consistency at any level.

#### Validation with Epidemic Potential Predictions

The goodness-of-fit tests do not identify any specific model, or even a class of model, as the best all-around fit to the data. A comparison of the epidemic potential predicted by each model offers a different perspective. The predicted threshold transmissibility  $\tau_c$  and the 95% confidence interval around the threshold, gives us a direct estimate of the probability of an epidemic for any

given level of integrated transmissibility  $\bar{\tau}$ . We know that there is no generalized HIV/AIDS epidemic in the United States,<sup>35</sup> while the prevalence of other STI varies substantially: from under 1% for gonorrhea,<sup>36</sup> and about 4% for chlamydia,<sup>36</sup> to over 20% for human papilloma virus (HPV)<sup>37</sup> and herpes simplex virus-2 (HSV-2).<sup>38</sup> This allows us to evaluate our estimates of  $\tau_c$  in 2 ways. First, if the predicted transmissibility threshold is zero such that any  $\bar{\tau}$  is likely to produce a generalized epidemic, we can conclude that the predictive model fails. Second, if the predicted transmissibility threshold is greater than zero, the potential for a generalized epidemic depends on  $\bar{\tau}$ . The value of  $\bar{\tau}$  varies by pathogen, so in principle, if the estimated  $\tau_c$  lies below  $\bar{\tau}$  for that pathogen, we would expect generalized spread of that pathogen in the United States, and if it lies above  $\bar{\tau}$ , we would not. Our estimates of  $\tau_c$  can be evaluated accordingly.

There are a variety of transmissibility estimates in the literature for different pathogens, although there is considerable uncertainty about most of these, and considerable variability in what they represent (e.g., per coital act vs. per partnership and varying time periods for the partnership estimates). For reference, we show a generally accepted estimate of  $\bar{\tau}_{\text{HIV}}$  here that is drawn from a study of discordant couples in Uganda, 0.162.<sup>28</sup> For a number of reasons reviewed in the discussion, this is likely to be an overestimate for the United States, but it is not out of line with older US estimates of  $\bar{\tau}_{\text{HIV}}$  based on transmission from transfusion recipients to their partners.<sup>39</sup> The analysis and presentation of the results allow the reader to make comparisons to alternative values.

The results can be seen in Figure 2. For the power law models (top panel) the predicted epidemic potential of these networks is infinite: the transmissibility thresholds are all zero and the 95% confidence intervals do not include any nonzero values. As with the naïve estimates based on linear regression, the more accurate estimates from these models predict generalized epidemics of all STIs in the United States. The predictions from the power law models simply do not fit the evidence.

Notes: The black line indicates no epidemic threshold exists; the grey line indicates the estimated network epidemic potential required to produce an epidemic of HIV, based on a published estimate of from Uganda.<sup>29</sup> The grey band by HIV transmissibility estimates from the United States.<sup>39</sup> NHLS-Cont

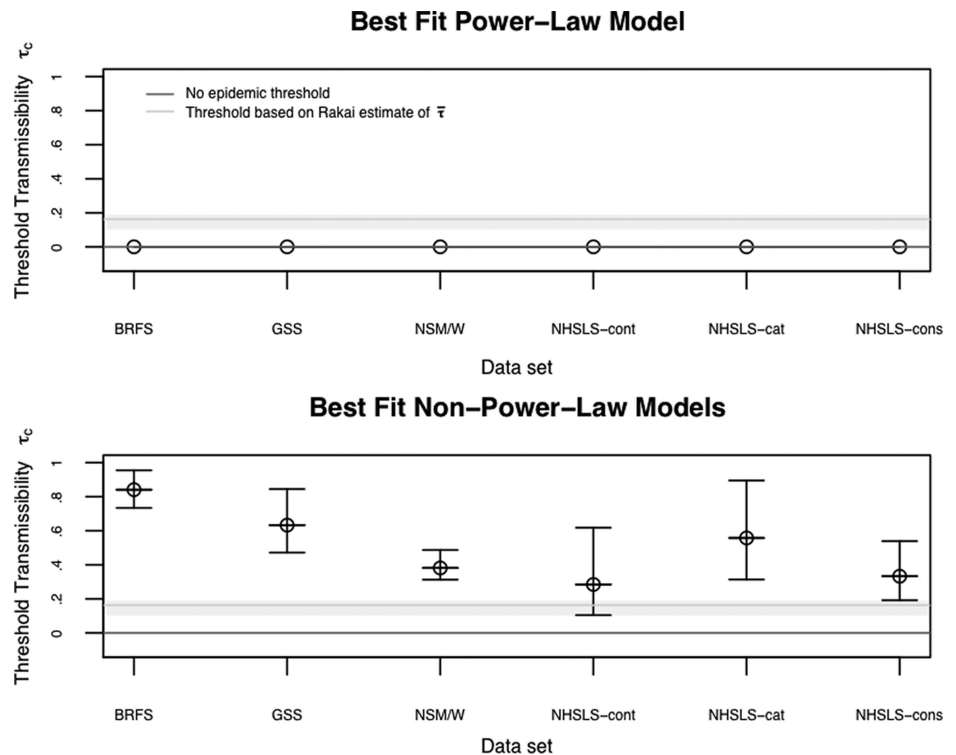


Fig. 2. The threshold transmissibility and 95% confidence intervals predicted for each of the 6 degree distributions by the best fit power-law model (top panel) and nonpower-law models (bottom panel).

is data from the continuous response question; -Cat is data from the categorical response question; -Cons is data from the constructed variable.

For the best fit nonpower-law models (bottom panel), the predicted epidemic potentials are much lower, and the predicted transmissibility thresholds correspondingly higher. All of the threshold transmissibility estimates are above 0.2 and 3 of the 6 are above 0.5. None of the 95% confidence intervals includes zero. As a result, these models predict that an epidemic threshold does exist, and the value of  $R_0$  will depend on the transmissibility  $\bar{\tau}$  of the STI. The implications for any specific STI can be read off of the graph directly: for  $R_0$  to be at least 1, the integrated transmissibility  $\bar{\tau}$  for that STI would have to be larger than the predicted threshold  $\tau_c$  shown on the vertical axis. For reference, we plot an estimate of  $\bar{\tau}_{\text{HIV}}$  drawn from a study of discordant couples in Uganda, 0.162,<sup>28</sup> in the grey line. All of the predicted thresholds under the best fitting nonpower-law models lie above this line, and only one of the confidence intervals includes it. Based on this estimate of  $\bar{\tau}_{\text{HIV}}$ , none of these models would predict a generalized epidemic of HIV in the United States. Alternative lines can be drawn for other values of transmissibility.

### Discussion

The results of this study are threefold. First, when tested against alternatives, power law models do not enjoy consistent support from the data. Using goodness of fit to the degree distributions as the criteria, several plausible social process models, representing either a search and stop rule mechanism or a 2-stage vetting model, are as likely to be the best fitting model or ranked in the top 3. These 2 model classes are flexible and have several very natural and intuitive interpretations. The search and stop rule models, for example, can be thought of as selecting new partners until some set of criteria are satisfied, for example, dating until a spouse is found

or dating until enough experience in the social world has been acquired. The 2-stage vetting models can be interpreted as meeting some number of people and selecting a partner from those acquaintances or selecting a partner from the people who live in the same area. The best overall fitting social process model was the Negative Binomial, which can be interpreted in 1 of 2 ways: Either people use a search and stop rule, or they acquire partners at a fixed, but heterogeneous, rate.

Second, using the prediction of the epidemic threshold in the United States as the criteria, all of the power law models fail completely: All predict a generalized epidemic of HIV—and of every other sexually transmitted infection—based on every data set examined here. It is worth emphasizing that the estimates of the power law models generate this prediction regardless of the level of pathogen transmissibility: Any pathogen with a nonzero probability of transmission is expected to cause an epidemic.

All of the best-fitting social process models, by contrast, predict epidemics only if transmissibility is sufficiently high, on the order of 0.3 to 0.6, with 1 outlier at 0.8. The implication is that the epidemic potential in this network is probably too low to sustain a generalized epidemic for STI with average integrated transmissibility per partnership below 20% to 30%. HIV is thought to have a transmissibility in this range, so the social process model estimates are consistent with the observed low prevalence of HIV in the United States. By contrast, the HPV is estimated to have a per partnership transmissibility of about 60%,<sup>40</sup> and HPV prevalence in the United States is about 25%. So here too, the social process model estimates are consistent with the data.

Many (if not all) of the curable STI (e.g., chlamydia, gonorrhea, syphilis, trichomoniasis) are estimated to have similarly high transmissibilities—50% to 70% per partnership.<sup>41–45</sup> As the social process models would predict, all these STIs are also endemic in the United States, though at relatively low prevalences, ranging from 0.05% for syphilis<sup>46</sup> to 4% for chlamydia.<sup>36,47</sup> The low



observed prevalence, given the very high transmissibility, may seem inconsistent with our estimate of the epidemic potential, but lower prevalence would be expected given the shorter duration of infection. For these STI, it may also be more appropriate to base the estimate of  $\tau_c$  on partners in the last month rather than the last year. This would lead to substantially lower estimates of epidemic potential and be more consistent with the observed prevalence.

In the main, the social process model estimates seem to be consistent with what we know about STI transmissibilities and prevalence in the United States. The only exception we can find is HSV-2. HSV-2 is thought to have a low transmissibility similar to HIV,<sup>48</sup> yet the prevalence of HSV-2 in the United States is over 20%.<sup>38</sup> So, that is a bit of a puzzle. It is not simply a problem with the social process models, however, as this is a pattern that is not consistent with the basic transmission dynamic modeling framework.

The complete failure of the scale-free models to make reasonable epidemic predictions is striking, particularly in light of the very strong public policy relevance claimed for these models. For example,<sup>7</sup> write that their "... most important finding is the scale-free nature of the connectivity of an objectively defined nonprofessional social network." The public health implications, they explain, are that:

"... the measures adopted to contain or stop the propagation of diseases in a network need to be radically different for scale-free networks. Single-scale networks are not susceptible to attack at even the most connected nodes, whereas scale-free networks are resilient to random failure but are highly susceptible to destruction of the most connected nodes."

It is claims like these on the deep relevance of the previous scale free estimates for STD prevention that we criticize here. Other examples of such claims are given in Ref. 6.

The third, and perhaps more important finding, is that none of the models tested here consistently provide the best fit to the data. Model rankings varied by survey, by measure within survey, by sex, and by the upper-tail cutoff used to restrict the data used to fit the models. One interpretation of this finding is that none of these models does a good job capturing the underlying mechanism that generates sexual partnership networks. The large confidence intervals around the predicted thresholds, even by the best fit models, illustrate just how little leverage degree distributions alone provide in establishing the likelihood of an epidemic.

A good empirical example of the predictive limitations of the degree-based model is the pervasive and long standing racial disparity in STI prevalences. Among non-Hispanic blacks in the United States, the relative risk of infection ranges from about 6 for chlamydia<sup>36</sup> to over 20 for gonorrhea<sup>36</sup> and HIV.<sup>35</sup> These disparities are not explained by group differences in rates of partnership acquisition<sup>49</sup> so would not be produced by any of the simple degree-based heterogeneity models considered here. One of the most important features of STI prevalence in the United States is thus beyond the explanatory capacity of these models.

To inform and guide public health, network-based approaches to epidemic modeling must take in to account the key structural features of networks that influence transmission, and the behaviors that generate these structures. This kind of research is being done, albeit with less publicity. An example is the work on "assortative mixing," the tendency for people to choose partners like themselves in terms of age,<sup>50</sup> race/ethnicity,<sup>51,52</sup> sexual role,<sup>53</sup> and possibly activity level.<sup>54</sup> Assortative mixing can lead to patchy clusters in the network with little connection between them, and it is possible to obtain analytic solutions for the effect of such mixing on epidemic thresholds.<sup>55,56</sup> Another important form of heterogeneity is in the timing and sequence of partnerships. This includes

the impact of variations in monogamous partnership duration<sup>57</sup> and the prevalence of concurrent partnerships.<sup>58</sup> Lon-term serial monogamy can dramatically reduce epidemic potential, but even small amounts of concurrency can reestablish connectivity and substantially increase the potential for spread. A good review of these patterns and their impacts on transmission can be found in Ref. 59. The implication of this body of work is not that the simpler models should be abandoned, just that they should be properly evaluated and their limitations properly communicated.

The public health implications of our results are clear: The radical implications of the "scale-free" models are theoretically intriguing, but in all likelihood empirically irrelevant for generalized HIV epidemics. Broad population-based prevention strategies such as education programs risk behavior reduction efforts, condom distribution, and antiretroviral therapies may therefore be effective in reducing the spread of HIV and other incurable STI and should not be abandoned in favor of exclusive targeting of high-degree hubs.

## References

- Hethcote HW, Yorke JA. *Gonorrhea Transmission Dynamics and control*. Berlin, Germany: Springer Verlag, 1984.
- Thomas JC, Tucker MJ. The development and use of the concept of a sexually transmitted disease core. *J Infect Dis* 1996; 174(Suppl 2): S134–S143.
- UNAIDS/WHO. *Second Generation Surveillance for HIV/AIDS: The Next Decade*. Geneva, Switzerland: UNAIDS, 1999.
- Morris M. Network epidemiology: A handbook for survey design and data collection. In: *International Studies in Demography*. Oxford, UK: Oxford University Press, 2004.
- Handcock M, Jones J. Interval estimates for epidemic thresholds in two-sex network models. *Theor Popul Biol* 2006; 70:125–134.
- Handcock MS, Jones JH, Morris M. *On Sexual Contacts and Epidemic Thresholds, Models and Inference for Sexual Partnership Distributions*. Seattle, WA: University of Washington Center for Statistics and the Social Sciences, 2003.
- Liljeros F, Edling CR, Amaral LAN, et al. The web of human sexual contacts. *Nature* 2001; 411:907–908.
- Schneeberger A, Mercer CH, Gregson SA, et al. Scale-free networks and sex transmitted diseases: A description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sex Transm Dis* 2004; 31:380–387.
- Laumann EO, Gagnon JH, Michael RT, et al. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago, IL: University of Chicago Press, 1994.
- Lewin B. *Sex in Sweden*. Stockholm: National Institute of Public Health, 1996.
- Hubert M, Bajos N, Sandfort T. *Sexual Behavior and HIV/AIDS in Europe: Comparisons of National Surveys*. London, UK: UCL Press, 1998.
- Aral SO. Sexual network patterns as determinants of STD rates: Paradigm shift in the behavioral epidemiology of STDs made visible. *Sex Transm Dis* 1999; 26:262–264.
- Youm Y, Laumann E. Social network effects on the transmission of sex transmitted diseases. *Sex Transm Dis* 2002; 29:689–697.
- Davis JA, Smith TW, Peter V, Marsden PV. *General Social Surveys, 1972–2002: Cumulative Codebook*. Chicago, IL: NORC, 2003.
- Tanfer K. National survey of men—Design and execution. *Fam Plann Perspect* 1993; 25:83–86.
- Tanfer K. A user's guide to the machine-readable files and documentation. 1991 national survey of women. Los Altos, CA: Sociometrics Corporation; 1995.
- Denzo Z, Barabasi A-L. Halting viruses in scale-free networks. *Physical Review E Stat Nonlin Soft Matter Phys* 2002; 65.
- Amaral LAN, Scala A, Barthelemy M, et al. Classes of Small-World Networks. *Proc Natl Acad Sci U S A* 2000; 10:11149–11152.
- Lloyd AL, May RM. Epidemiology. How viruses spread among computers and people. *Science* 2001; 292:1316–1317.

20. Huttenlocher J, Hedges LV, Bradburn NM. Reports of elapsed time—Bounding and rounding processes in estimation. *J Exp Psychol Learn Mem Cogn* 1990; 16:196–213.
21. Morris M. Telling tails explain the discrepancy in sexual partner reports. *Nature* 1993; 365:437–440.
22. Handcock MS, Jones JH. Likelihood-based inference for stochastic models of sexual network formation. *Theor Popul Biol* 2004; 65: 413–442.
23. Simon HA. On a class of skew distribution functions. *Biometrika* 1955; 42:34–48.
24. Irwin J. The place of mathematics in medical and biological statistics. *J R Stat Soc Ser A* 1963; 126:1–45.
25. Barabasi A-L, Albert R. Emergence of scaling in random networks. *Science* 1999; 286:509–512.
26. Albert R, Barabasi A-L. Topology of evolving networks: Local events and universality. *Phys Rev Lett* 2000; 85:5234–5237.
27. Hamilton DT, Morris M. Consistency of Self-Reported Sexual Behavior in Surveys. Seattle, WA: University of Washington Center for Studies in Demography and Ecology; 2007.
28. Gray RH, Wawer M, Brookmeyer R, et al. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. *Lancet* 2001; 357:1149–1153.
29. Laumann EO, Gagnon JH, Michael RT, et al. National health and social life survey, 1992: United States (computer file). Ann Arbor, MI. Inter-university Consortium for Political and Social Research (distributor), 1995.
30. Perline R. Strong, weak and false inverse power laws. *Stat Sci* 2005; 20:68–88.
31. Casella G, Berger R. *Statistical Inference*, 2nd edn. Pacific Grove, CA: Duxbury, 2002.
32. Akaike H. A new look at the statistical model identification. *IEEE Trans Automatic Control* 1974; 19:716–722.
33. Burnham K, Anderson D. *Model Selection and Inference: A Practical Information-Theoretic Approach*, 2nd edn. New York: Springer, 2002.
34. Raftery AE. Bayesian model selection in social research. *Sociol Methodol* 1995; 25:111–163.
35. Morris M, Handcock MS, Miller WC, et al. Prevalence of HIV infection among young adults in the US: Results from the ADD Health Study. *Am J Public Health* 2006; 96:1091–1097.
36. Miller WC, Ford CA, Morris M, et al. Prevalence of chlamydial and gonococcal infection among young adults in the United States. *JAMA* 2004; 291:2229–2236.
37. Dunne EF, Unger ER, Sternberg M, et al. Prevalence of HPV infection among females in the United States. *JAMA* 2007; 297:813–819.
38. Fleming DT, McQuillan GM, Johnson RE, et al. Herpes simplex virus type 2 in the United States, 1976 to 1994. *N Engl J Med* 1997; 337:1105–1111.
39. Peterman T, Stoneburner R, Allen J, et al. Risk of HIV transmission from heterosexual adults with transfusion associated infections. *JAMA* 1988; 259:55–58.
40. Barnabas RV, Laukkanen P, Koskela P, et al. Epidemiology of HPV 16 and cervical cancer in Finland and the potential impact of vaccination: Mathematical modelling analyses. *PLoS Med* 2006; 3:624–632.
41. Lin J-SL, Donegan SP, Heeren TC, et al. Transmission of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* among men with urethritis and their female sex partners. *J Infect Dis* 1998; 178:1707–1712.
42. Garnett GP, Aral SO, Hoyle DV, et al. The natural history of syphilis: Implications for the transmission dynamics and control of infection. *Sex Transm Dis* 1997; 24:185–200.
43. Hobbs MM, Lapple DM, Lawing LF, et al. Methods for detection of *Trichomonas vaginalis* in the male partners of infected women: Implications for control of trichomoniasis. *J Clin Microbiol* 2006; 44:3994–3999.
44. Katz BP, Caine JR. Estimation of transmission probabilities for chlamydial infection. In: *International Symposium on Human Chlamydial Infections*. New York: Cambridge University Press, 1990:567–570.
45. Hooper RR, Reynolds GH, Jones OG, et al. Cohort study of venereal disease: The risk of gonorrhea transmission from infected women to men. *Am J Epidemiol* 1978; 108:136–144.
46. CDC. Sexually Transmitted Disease Surveillance 2005 Supplement Syphilis Surveillance Report. Atlanta, GA: Department of Health and Human Services Centers for Disease Control and Prevention, 2006.
47. Miller WC, Swygard H, Hobbs MM, et al. The prevalence of trichomoniasis in young adults in the United States. *Sex Transm Dis* 2005; 32:593–598.
48. Wald A, Langenberg AGM, Link K, et al. Effect of condoms on reducing the transmission of herpes simplex virus type 2 from men to women. *JAMA* 2001; 285:3100–3106.
49. Hallfors DD, Iritani BJ, Miller WC, et al. Sexual and drug behavior patterns and HIV and STD racial disparities: The need for new directions. *Am J Public Health* 2007; 97:125–132.
50. Gregson S, Nyamukapa CA, Garnett GP, et al. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet* 2002; 359:1896–1903.
51. Morris M. Data driven network models for the spread of infectious disease. In: Mollison D, ed. *Epidemic Models: Their Structure and Relation to Data*. Cambridge, UK: Cambridge University Press, 1995.
52. Laumann E, Youm Y. Racial/ethnic group differences in the prevalence of sexually transmitted diseases in the United States: A network explanation. *Sex Transm Dis* 1999; 26:250–261.
53. Goodreau SM, Goicochea LP, Sanchez J. Sexual role and transmission of HIV type 1 among men who have sex with men, in Peru. *J Infect Dis* 2005; 191:s147–s158.
54. Garnett GP, Anderson RM. Sexually transmitted diseases and sexual behavior: Insights from mathematical models. *J Infect Dis* 1996; 174(suppl 2):s150–s161.
55. Newman MEJ. Assortative mixing in networks. *Phys Rev Lett* 2002; 89:208701.
56. Fulford G, Roberts MG, Heesterbeek JAP. The metapopulation dynamics of an infectious disease: Tuberculosis in possums. *Theor Popul Biol* 2002; 61:15–29.
57. Haderl K, Waldstatter R, Worz B. Models for pair formation in bisexual populations. *J Math Biol* 1988; 26:635–649.
58. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. *AIDS* 1997; 11:641–648.
59. Morris M, Goodreau S, Moody J. Sexual networks, concurrency, and STD/HIV. In: Holmes KK, Sparling PF, Stamm WE, et al., eds. *Sexually Transmitted Diseases*. 4th edn. New York: McGraw-Hill, 2007.

Appendix: Model Fit to the Degree Distribution of the US Population Age 20–39 Based on BIC

Cutoff	Males				Females			
	Model	Log-lik	AICC	BIC	Model	Log-lik	AICC	BIC
Data: BRFS								
1	NBY	-20426	40,860	40,891	Waring	-17765	35,536	35,560
	NB	-20462	40,929	40,953	DP	-17775	35,554	35,570
	Waring	-20474	40,953	40,977	NBY	-17767	35,542	35,574
2	PLN	-20367	40,742	40,773	DP	-17765	35,536	35,560
	Waring	-20399	40,806	40,838	Yule	-17769	35,545	35,569
	NBY	-20412	40,834	40,873	DPE	-17765	35,537	35,570
3	PLN	-20367	40,744	40,783	DP	-17763	35,535	35,567
	Yule	-20392	40,791	40,823	Yule	-17767	35,542	35,575
	DP	-20396	40,801	40,832	DPE	-17763	35,537	35,577
Data: GSS								
1	NB	-3700	7,407	7,425	Waring	-3364	6,733	6,752
	DP	-3842	7,687	7,699	DP	-3370	6,743	6,756
	Waring	-3838	7,682	7,700	Yule	-3370	6,744	6,757
2	NB	-3744	7,495	7,519	Yule	-3363	6,732	6,751
	Geo	-3753	7,512	7,530	DP	-3366	6,738	6,757
	DPE	-3823	7,653	7,677	PLN	-3362	6,732	6,757
3	NB	-3736	7,483	7,513	Yule	-3363	6,734	6,759
	Geo	-3766	7,540	7,564	DP	-3363	6,735	6,760
	DP	-3809	7,625	7,649	PLN	-3362	6,734	6,765
Data: NSM/NSW								
1	NBY	-4496	9,000	9,024	NB	-1931	3,868	3,884
	NB	-4510	9,027	9,045	Yule	-1935	3,874	3,885
	Waring	-4513	9,033	9,051	DPE	-1931	3,869	3,885
2	PLN	-4474	8,955	8,980	Geo	-1931	3,868	3,884
	Waring	-4483	8,973	8,998	Yule	-1934	3,874	3,890
	DPE	-4493	8,994	9,018	PLN	-1930	3,868	3,890
3	Yule	-4476	8,960	8,984	DP	-1928	3,863	3,885
	DP	-4477	8,962	8,987	Yule	-1928	3,864	3,886
	PLN	-4473	8,957	8,987	NB	-1925	3,861	3,888
Data: NHLSL, 1 open-ended question								
1	DP	-1716	3,436	3,446	NB	-1545	3,097	3,113
	Waring	-1716	3,437	3,453	Waring	-1554	3,115	3,131
	DPE	-1716	3,438	3,454	NBY	-1551	3,110	3,132
2	Yule	-1715	3,436	3,452	Geo	-1544	3,093	3,110
	DP	-1716	3,438	3,453	NB	-1544	3,095	3,117
	Waring	-1715	3,438	3,459	DPE	-1544	3,095	3,117
3	DP	-1715	3,437	3,458	Geo	-1543	3,095	3,117
	Yule	-1715	3,438	3,459	Yule	-1547	3,101	3,123
	Waring	-1715	3,439	3,465	DPE	-1543	3,097	3,124
Data: NHLSL, 1 Categorical question								
1	NB	-1602	3,210	3,225	DP	-1613	3,230	3,241
	DP	-1636	3,276	3,286	Waring	-1611	3,229	3,245
	Waring	-1634	3,275	3,291	NBY	-1613	3,234	3,256
2	Geo	-1610	3,225	3,241	DP	-1611	3,228	3,245
	NB	-1608	3,223	3,244	Yule	-1613	3,232	3,249
	Yule	-1633	3,272	3,288	Waring	-1611	3,231	3,253
3	Geo	-1611	3,230	3,251	DP	-1611	3,230	3,252
	NB	-1610	3,229	3,255	Yule	-1613	3,234	3,256
	DP	-1631	3,270	3,291	Waring	-1611	3,231	3,259
Data: NHLSL, Constructed by Laumann et al., 1995								
1	NB	-1653	3,313	3,328	DP	-1625	3,253	3,264
	NBY	-1681	3,370	3,391	Waring	-1621	3,248	3,265
	DP	-1697	3,397	3,408	NBY	-1620	3,248	3,270
2	NB	-1656	3,320	3,341	Yule	-1619	3,245	3,261
	PLN	-1680	3,367	3,388	DP	-1621	3,247	3,264
	Waring	-1680	3,368	3,389	Waring	-1619	3,247	3,269
3	Geo	-1658	3,324	3,345	DP	-1618	3,245	3,267
	NB	-1658	3,326	3,352	Yule	-1619	3,246	3,268
	PLN	-1677	3,364	3,390	Waring	-1618	3,247	3,274