On "Sexual contacts and epidemic thresholds," models and inference for Sexual partnership distributions

Mark S. Handcock[1]
Departments of Statistics and Sociology
Center for Statistics and the Social Sciences
Department of Statistics
University of Washington, Seattle, WA

James Holland Jones
Center for Studies in Demography and Ecology
Center for Statistics and the Social Sciences
Center for AIDS and Sexually Transmitted Diseases
University of Washington, Seattle, WA

Martina Morris
Departments of Sociology and Statistics
Center for Studies in Demography and Ecology
University of Washington, Seattle, WA

[1]Corresponding Author: Center for Statistics and the Social Sciences, Box 354320, Seattle, WA 98195-4320, email: `handcock@stat.washington.edu`, Phone: 206-221-6874, Fax: 206-221-6873

Abstract

Recent work has focused attention on statistical inference for the population distribution of the number of sexual partners based on survey data.

The characteristics of these distributions are of interest as components of mathematical models for the transmission dynamics of sexually-transmitted diseases (STDs). Such information can be used both to calibrate theoretical models, to make predictions for real populations, and as a tool for guiding public health policy.

Our previous work on this subject has developed likelihood-based statistical methods for inference that allow for low-dimensional, semi-parametric models. Inference has been based on several proposed stochastic process models for the formation of sexual partnership networks. We have also developed model selection criteria to choose between competing models, and assessed the fit of different models to three populations: Uganda, Sweden, and the USA. Throughout this work, we have emphasized the correct assessment of the uncertainty of the estimates based on the data analyzed. We have also widened the question of interest to the limitations of inferences from such data, and the utility of degree-based epidemiological models more generally.

In this paper we address further statistical issues that are important in this area, and a number of confusions that have arisen in interpreting our work. In particular, we consider the use of cumulative lifetime partner distributions, heaping and other issues raised by Liljeros *et al.* in a recent working paper.

1 Introduction

In two recent papers (Jones and Handcock, 2003a,b), we have been critical of the statistical methodology used to asses the scaling behavior of sexual partnership distributions. Because of the important public health implications of research into the properties of sexual contact networks, we argue that tools used to evaluate these properties must be of the highest scientific quality. Two points in particular are fundamental to rigorous science which underlies policy-making:

1. An accurate accounting of uncertainty, since a large segment of sound public policy (including public health policy) is the management of risk under uncertainty (Tanur et al., 1989; Berger, 1986; Tuljapurkar et al., 2000).

2. The transparent application of methodology, including precise definitions of key terms, and the specification of statistical and mathematical tools and the assumptions which underlie them.

Liljeros et al. (2001, 907) write that their "...most important finding is the scale-free nature of the connectivity of an objectively-defined non-professional social network." Though the term "scale-free" is never defined, the public health implications of this finding are articulated in the following paragraph:

> ...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 (Liljeros et al., 2001, 907).

By "destruction of most connected nodes," we assume they mean treating (or sequestering?) highly promiscuous people.

Targeted interventions, of course, are nothing new in public health, as we note in Jones and Handcock (2003a), citing work targeting populations such as commercial sex workers (Ford and Koetsawang, 1999), clients of sex workers (Morris, 2001), army recruits (Nelson et al., 2002), persons with concurrent partners (Morris et al., 1997), or injection drug users (Neaigus, 1999), which have a proven record in reducing disease incidence.

In addition to targeted interventions, standard practice in infectious disease epidemiology attempts to eradicate infectious agents by vaccinating (or treating) a critical fraction of the population (see e.g., Anderson and May, 1991).

One clear implication of the "radically different" interventions required by scale-free networks is that the critical vaccination fraction is unity – all members of the population must be *successfully* vaccinated to eradicate the disease. These sentiments are reflected in a commentary which accompanied the piece by Liljeros et al. (2001) by Philip Ball: "Sexually transmitted diseases (STDs) call for discrimination. Containing the spread of an STD by focusing on promiscuous individuals, who are most likely to pass it on, should be cheaper and more effective than large-scale random campaigns, according to two new mathematical analyses."

Similarly, in a popular piece on scale-free networks written by science writer David Cohen for the *New Scientist* (13 April 2002, 2338:24-29), the implications are laid out starkly: "It is common sense that a programme of vaccination against sexually-transmitted diseases should try to reach the most promiscuous individuals first. But the idea that health campaigns may be utterly worthless if they miss these people is a shock." In this same article, Frederik Liljeros is quoted as saying that "We can attempt to stop the spread of a virus by blindly vaccinating huge groups, but without treating these key individuals we may never bring it under control."

Economic arguments are clearly central to the public health implications here (Dezső and Barabási, 2002, 1):

> However, due to economic or policy considerations the number of available cures is often limited. This applies to AIDS, for which relatively effective but prohibitively expensive cures are available, unable to reach the most affected segments of population due to economic considerations.

We are concerned that such economic arguments may imbue the scale-free network idea with an unwarranted attractiveness for some policy makers. Specifically, "destroying hubs" may be seen as a cheap fix to the problem of the global HIV/AIDS pandemic and jeopardize funding for capital-intensive but proven health interventions such as safe sex campaigns (which are always on politically shaky ground), vaccine development, the prevention of mother-to-child transmission, and the treatment of the destitute poor and other people in resource-poor settings of the developing world (Farmer, 1999, 2003; Kim et al., 2000).

A primary purpose of our recent papers on inference for the degree distributions of sexual contact networks was to demonstrate that the empirical support for the belief that transmission-reduction interventions cannot eradicate STDs is very weak (Jones and Handcock, 2003a,b; Handcock and Jones, 2003b,a). We show that, even with the small samples that are available one can say with high statistical confidence that the estimates are outside the infinite variance range, and hence give additional rationale for the existence of thresholds. The second purpose was to promote the use of sound statistical methodology to infer the models from data. This will be an active area of research.

Our recent work (i.e., Handcock and Jones (2003b,a); Jones and Handcock (2003a,b)) has elicited a response from the authors of a prior analysis of the Swedish dataset (Liljeros et al., 2001).[1] These authors interpreted our comments as critical, and their response has been defensive (e.g., Liljeros et al., 2003b,a). However, we were pleased to see an improvement in their methods based on an understanding of our work. However, confusions remain and more were introduced by their response. In the next section we review some of these. Our comments are not meant as criticisms of Liljeros *et al.*, but to clarify the terms and issues they raise. We have found the debate both illuminating and useful for shaping our thinking on the issues, and here capitalize on our ability to write in the more discursive format allowed by a working paper to address, point-by-point some of the criticisms and points that we see as confused.

---

[1]We should note that this work takes the form of three working papers at the Center for Statistics and the Social Sciences (working paper #'s 21, 23, and 29). In Liljeros et al. (2003b), the authors cite the title of wp29, the number of wp23, and appear to be responding, at least in part, to wp21. Working paper 23 corresponds to Jones and Handcock (2003a). There is a second, brief paper in press at *Nature*, and wp29 has been submitted for publication.

2 Models, Issues, Definitions and Interpretations

Unless otherwise noted, the following quotes are from Liljeros et al. (2003b).

1. "Specifically, Jones and Handcock analyze the probability distribution function [sic] $p(k)$, a function from which one obtains less reliable information than from the cumulative distribution,"

   (a) First a clarification on nomenclature: The function $p(k)$ has discrete support and is usually referred to as the probability mass function (PMF), a convention we follow in our papers. The term probability distribution function is often used as a synonym for the cumulative distribution function. It is possible that there is confusion with the probability density function with is used for continuous variables. It seems clear from the context that Liljeros *et al.* mean the PMF.

   (b) In our work, we model the probability distribution which is completely characterized by the PMF or, equivalently, the cumulative distribution function. As both can be mathematically derived from the other in this case, it makes no sense to say that we are modeling one rather than the other. It is more accurate to say we are modeling the distribution.

   (c) A clear distinction needs to be drawn between the underlying probabilistic model and the various statistics used to estimate the model. Specifically, between the PMF and CDF on one hand and the empirical PMF, CDF or survival function on the other. The former are of primary interest and the latter are mainly useful to infer the former. So there may be differing information content in the statistics employed for the same model. An advantage of likelihood-based methods is they encapsulate the information in the data via sufficient statistics for the model.

   (d) In light of these comments, the claim that the PDF has less reliable information in it than the CDF is nonsensical. Liljeros *et al.* may be confusing issues of statistical estimation with those of probabilistic representation. We note that the information we use represent sufficient statistics for the model, while the implicit method that Liljeros *et al.* use for their estimation procedure makes it unclear what information is actually used. This means the quantification of the uncertainty of their estimate is very difficult. As we discuss in our papers, the estimates they provide of the uncertainty are based on an inappropriate method and will, in general, be inaccurate. Note also that the subjective aspects of the curve fitting method hinder reproducibility by others or on similar datasets. For example, should we use a cutoff of 4, 5, 19 or 20? Should we fit the line to all points in the empirical CDF or only those where there are jumps? Each of these require judgment that is not reflected in the interval estimates.

2. "In the following, we argue that the claims of Jones and Handcock can be interpreted in a misleading fashion.

   Hopefully not. We urge readers not to misinterpret our claims.

3. Liljeros *et al.* correctly note that the existence of epidemic thresholds depends on the structure of the network and other factors beside the distribution of the number of partners in the last year. We agree with these statements and are gratified to see Liljeros *et al.* repeat them(Handcock and Jones, 2003b). Our work suggests that the evidence from the distribution of the number of partners in the last year does not support infinite variance. The jump from inference to epidemic thresholds requires a model for the complex structure of social networks, of which the distribution of the number of partners is only one component. We show in Handcock and Jones (2003a) that under one state-of-the art model (viz., Newman (2002)), these thresholds exist.

   Liljeros *et al.* also note that mathematical models are approximations to the real-world network. Many have been tempted to extrapolate from such approximation representations to the real-world without careful assessment of the quality of the approximation. This means that one answer to the question of finite variance for real-world sexual networks is, trivially, yes! However, this is an age old debate: some have argued the much of the mathematical research on networks is not relevant to epidemiology and public health because of the difficulty and rarity with which justification for the jump from mathematical models to the real-word behavior is made. A common complaint is that the models are too simple (as here) or the mathematics too complicated (Brauer and van den Driessche, 2002). However, this is clearly too extreme as much has been learned from both their qualitative and quantitative behavior (Anderson and May, 1991; Brauer and van den Driessche, 2002). Much more can potentially be learned, but only if sufficient justification for the mathematical approximations is made. We believe that energies are best spent addressing this fundamental research question rather than making radical and weakly supported claims.

4. "The estimates by Jones and Handcock in fact agree with the estimates already reported by us"

   In one of our papers we tried to privilege the simplest proposed power-law model (Albert and Barabási, 2000) by only comparing models with different scaling exponents (Jones and Handcock, 2003b). Based on likelihood-based estimates and the Bayesian Information Criterion (BIC), our model selection procedure found the best fitting Yule model for both males and females to be for degrees exceeding k_{min}=1. Liljeros et al. (2001) used $k_{min} = 5$ for men (retaining 2% of the data), and $k_{min} = 4$ for women without accounting for this in their quantification of uncertainty. These are different models. This statement can be seen visually in Figure 1 displaying the range of interval estimates of the scaling parameter for the probability mass function for Swedish males and females (Jones and Handcock, 2003b). Comparisons of the intervals as k_{min} increases shows the sensitivity of the 95% bootstrap confidence intervals to the upper tail of the partnership distribution, defined as $k > k_{min}$, where k is the number of sex partners in the previous 12 months. For both males and females the best-fitting model has $k_{min} = 1$. By contrast, the intervals reported by Liljeros et al. (2001) are marked in red. For these values, our estimates of parameter uncertainty are six times greater than their reported intervals.
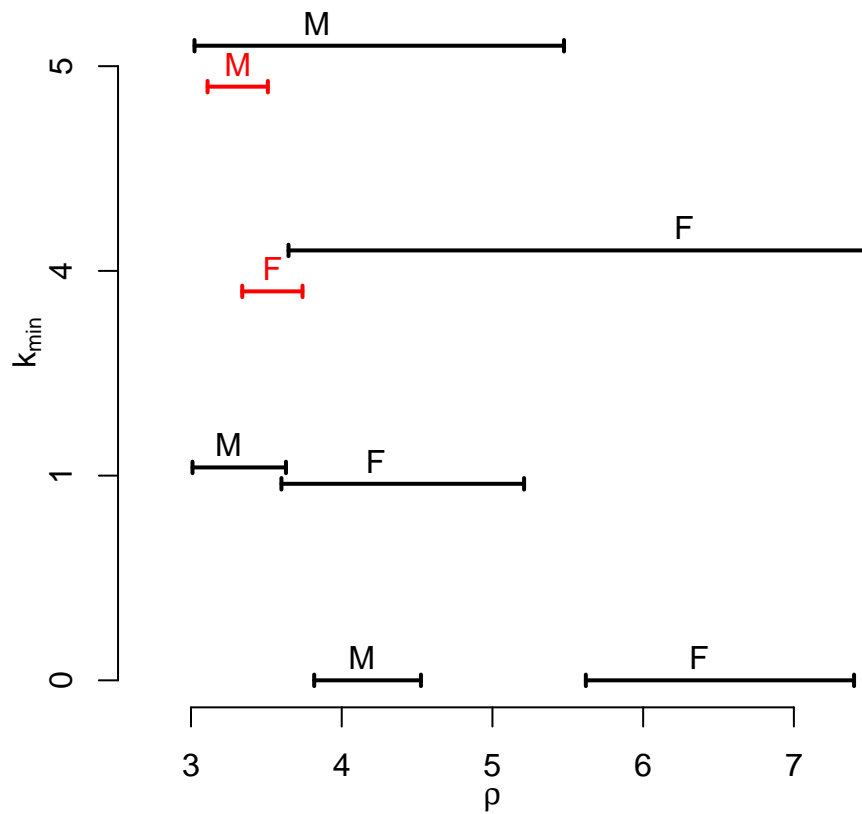
Figure 1: Interval estimates of the scaling parameter, ρ, for the generalized Yule probability mass functions for Swedish males and females. Data show the sensitivity of the 95% bootstrap confidence intervals to the upper tail of the partnership distribution, defined as $k > k_{min}$, where k is the number of sex partners in the previous 12 months. The intervals reported by Liljeros et al. (2001) are marked in red.

From a scientific perspective, these are not the same estimates.

We note that we also show (Jones and Handcock, 2003b; Handcock and Jones, 2003b; Jones and Handcock, 2003a) that the distributions are outside the infinite variance range for populations in two other countries (Uganda and the USA). This appears to triple the total number reported in the literature.

5. "Since $P(k)$[sic] is a summation of $p(k)$, the exponent ρ is related to the exponent α by the equation $1 + \alpha = \rho$. Hence the values of α that are somewhat above 2 reported in Ref. 2 do not differ from the values of ρ that are somewhat above 3 reported in Ref 1.

   As noted above, our interval estimates differ from theirs. The differences between the estimates is not a function of this simple relationship, nor do our papers confuse α and ρ. We have previously communicated this to Liljeros *et al.*.

6. "the Yule distribution does not display a power law for small k (Fig.1)."

   (a) Networks exhibiting power-law scaling have been referred to as "scale-free" networks in Amaral et al. (2000) and subsequent publications. This attribution is associated with properties of the implicit underlying stochastic mechanism, and the literature has been quite vague on the definitions of key terms. We have tried to be precise. Let K be the degree of a randomly sampled person from the population and $P(K = k)$ be the corresponding PMF. Let f and g be two functions with support the whole numbers. We take $f(k) \asymp g(k)$ to mean that there exist constants c_1, c_2 such that $0 \leq c_1 < f(k)/g(k) \leq c_2 < \infty$ for $k = 1, \ldots$. We then say that $P(K = k)$ has *power-law behavior* if $P(K = k) \asymp k^{-\rho}$. Such distributions are approximately *scale-free* in the sense that $P(K = k/c) \asymp P(K = k)$ for $c > 0$. Recent empirical work (Amaral et al., 2000; Liljeros et al., 2001) has claimed that some sexual network degree distributions have PMF for network degree of the form, $P(K = k) \approx k^{-\rho}, k >> 1$, where $P(K = k)$ is the probability of observing exact degree k and ρ is referred to as a scaling parameter.

   Under the definitions in our papers, the Yule distribution has power-law behavior. Indeed, power-law behavior is only defined asymptotically for k large.

   In Liljeros et al. (2001), Liljeros *et al.* state: "Scale-free networks, which are characterized by a power-law decay of the cumulative distribution $P(k) \approx k^{-\alpha}$, may be formed as a result of preferential attachment of new links between highly connected nodes [Simon 1955, Barabasi and Albert 1999]." Thus their use is consistent with our more formal one. The first reference is given to Simon, whom we cite as providing a stochastic mechanism resulting in the Yule distribution. Hence we struggle to make sense of their statement. It seems in their current paper that they use the term "power-law" to refer to the particular model (Johnson et al., 1992):

   $$P(K = k) = \frac{1}{\zeta(\rho)k^\rho}, \qquad\qquad k = 1, 2, \ldots \qquad (1)$$

   where $\zeta(\rho)$, the Riemann zeta function of ρ. We discuss this in detail (Jones and Handcock, 2003a; Handcock and Jones, 2003b). Linguists proposed an underlying

the stochastic mechanism called the *Zipf-Estroup law* (Estroup, 1916; Zipf, 1949) analogous to that for the Yule proposed by Simon. This is often referred to as the *discrete Pareto distribution* in analogy to the continuous version (which has been used as a model for degree distribution in May and Lloyd (2001)). It is also referred to as the *(Riemann) zeta distribution.* Two examples of this distribution are Zipf's law ($\rho = 1$) relating the relative frequency of words in a text to their rank, and Lotka's law ($\rho = 2$) for the number of authors making contributions.

In their current paper, power-law is not defined. However, it appears they are using the discrete Pareto in contrast to the Yule cited in their original paper.

We urge readers to use a consistent notation, with key terms defined. We have suggested a notation, but the reader can choose one appropriate for their circumstance.

(b) The choice among the discrete Pareto and the Yule should be based on the reasonableness of the stochastic mechanism that underlies them. They both exhibit power-law behavior. In Jones and Handcock (2003b); Handcock and Jones (2003b) we reference discussions of the different mechanisms.

(c) In their Table 1, Liljeros et al. (2003b) compare the discrete Pareto to the Yule fit for $k > 0$. We compared a range of models in Handcock and Jones (2003b) (which they cite). For $k > 0$ the shifted negative binomial model provides a better fit to the data than both the Yule and discrete Pareto (in terms of the corrected AIC and BIC). The negative binomial has finite variance, and implies an epidemic threshold exists. See Handcock and Jones (2003b) for details.

7. "As social scientists are no longer forced to select mathematical models on the basis of analytical tractability at the cost of realism [10], it is surprising that Jones and Handcock do not select from the more general models including preferential attachment published in the scientific literature [Albert and Barabasi 2002]."

This comment makes no sense. Our work (which they cite) is focused on the proposal, development and analysis of an array of degree models with underlying stochastic mechanisms including preferential attachment. These include models in Albert and Barabasi (2002)! In these papers we use likelihood-based model selection procedures to assess the fit of the different models to three large distributions of sexual partner counts: Uganda, Sweden, and the USA. Five of the six single-sex networks were fit best by the negative binomial model (not the Yule or other preferential attachment model). The American women's network was best fit by a power-law model, the Yule. For most networks, several competing models fit approximately equally well. These results suggest three conclusions: (1) no single unitary process clearly underlies the formation of these sexual networks, (2) behavioral heterogeneity plays an essential role in network structure, (3) substantial model uncertainty exists for sexual network degree distributions. Behavioral research focused on the mechanisms of partnership formation will play an essential role in specifying the best model for empirical degree distributions.

As far as we know, this is one of the first papers to appear that proposes different power-law and non-power-law models and assesses their empirical fit using principled

methods.

We urge Liljeros *et al.* to read the paper they cite. Our work has argued for the use of more realistic models - and evaluated candidates. We would have compared the models in this paper to models proposed by Liljeros *et al.*, but have not found any (except the single "analytically tractable" one they use in Liljeros et al. (2001)).

There are many forms of network structures that can be termed preferential attachment. The Yule and Waring are two examples induced by a dynamic mechanism where the probability that a tie is made with any particular individual is a function of that individual's degree (e.g., proportional). However, it is worthwhile separating causal processes from associative ones. If the Yule or Waring distribution fits the degree distribution it may not be that a pure preferential attachment mechanism is the cause. In fact, the apparent attachment is likely to be induced by assortative matching on partner characteristics and other local network characteristics (e.g., gender, race, age, economic status, martial status). In fact, disassortative mating is probably more common in the activity context (e.g., men with multiple partners, women monogamous), and the level will vary by culture/subgroup. (e.g., polygamy vs. bathhouse).

Finally - despite the attraction of physicists to preferential attachment models very few social scientists would recognize preferential attachment as a reasonable model for sexual behavior. It is, however, a reasonable model for technological and some other networks.

8. "Another difference between the analyses of Refs. 1 and 2 is that Jones and Handcock assume that the data are best described by a Yule distribution [Simon 1955]."

As noted in the previous item, the cited paper compared multiple models. Despite its simplicity the Yule has been proposed as a model for stochastic processes going back over 80 years. In Jones and Handcock (2003a,b) we assume, but did not *presume* the Yule model as was done in Liljeros et al. (2001). Ironically, this was done to show that even with the assumption of the model form, the empirical support for infinite variance was weak. Our analyses in the cited paper show that the Yule is not the best model. Table 1 in Liljeros et al. (2003b) verifies this.

3 The use of cumulative lifetime numbers of partners

The multiplicity of forms of networks provide great scope for analysis. Sexual partnerships and disease spread are co-evolving dynamic processes, and mathematical models are idealizations of complex social and behavioral dynamics. The very simplistic models considered here involve a severe reduction of that complexity.

Liljeros *et al.* argue for the modeling of survey information collected on cumulative lifetime number of sexual partners as input to models of sexual networks. We welcome this as such analysis has potential. However it exacerbates temporal aggregation of partnerships, and hence confounds the sequence of partnerships and concurrency. Both of these are important for the

spread of STDs(Morris et al., 1997). In addition to these conceptual issues, the statistical analysis is also fraught with difficulties, some of which we discuss below.

Given the grain of the data collection, the single year partner count comes closest to approximating the instantaneous network, the network that a pathogen must traverse. It will also give a more accurate measure of the STD incidence at the time of the survey. The use of the total partners will obscure the changes in the partnership networks over time. This is an essential consideration in Uganda which now has a mature or declining epidemic after years of rapid growth in incidence. Recent work by Ferguson and colleagues has clarified the connection between instantaneous and cumulative networks.

Further the survey data that Liljeros *et al.* consider does not have information on the total number of sexual partners - they only have the information on the number of partners acquired to this point in a person's life. Such 100% censoring greatly biases analyses that do not adjust for it. In particular, the exponents reported in the earlier paper are invalid. We chose the 1yr information because it does not suffer from this defect.

The relevance of cumulative "networks" generated from the cross-sectional survey used by Liljeros *et al.* is more problematic than they imply. They contain no information on the temporal patterning of partnerships. A person with degree 10 who contracts an STD at a time when he has one stable partner is not at risk of transmitting the pathogen to his previous nine partners.

If we are interested in the epidemiologically relevant networks for viral STDs, then static networks based on completely censored data require more justification than that supplied by Liljeros *et al.*. In particular, we welcome further justification from the authors or others for the use of this information.

As an aside, many STDs such as *Chlamydia trachomatis,* and other bacterial STIs would not be well modeled on a cumulative network (even if such networks were reasonable models of the risk structure of the population). Partner's age is clearly a risk factor for the non-fatal viral STDs such as HSV-2 or HPV. The period of infectivity of genital *C. trachomatis* in humans is not well understood, but estimates place it at around 12 months, with many individuals spontaneously clearing infection without treatment (Katz, 1992).

Newman (2002) gives the most advanced two-sex model for epidemic behavior based on partnership distributions. He shows that there will always be an epidemic (i.e., no transmission threshold) only if one of the distributions has an infinite variance. If both variances are finite then reducing transmissibility will reduce the epidemic threshold and even make an epidemic impossible. Newman, in his choice of data to use for his study of epidemic thresholds and transmissibility uses the 12 month partnership data, noting:

> "One should observe that the network studied in [Liljeros et. al. 2001] is a cumulative network of actual sexual contacts - it represents the sum of all contacts over a specified period of time. Although this is similar to other networks of sexual contacts studied previously [50, 51], it is not the network required by our models, which is the instantaneous network of connections (not contacts-see Section II). While the network measured may be a reasonable proxy for the network we need, it is not known if this is the case."

Our view is close to Newman on this - all network data provides information of interest, some are better than others, and it depends on the setting. The choice is of particular relevance
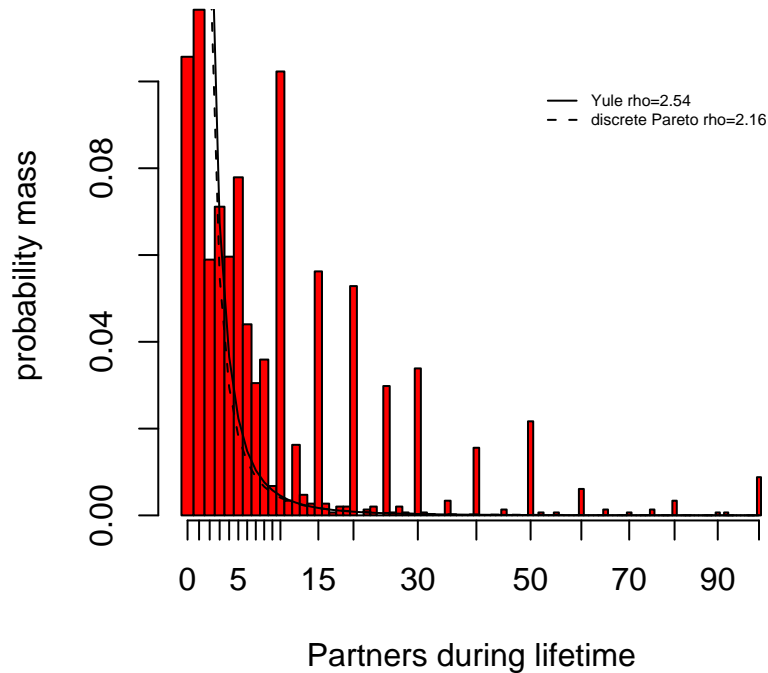
Figure 2: Reported cumulative lifetime degree distribution for Swedish men. The plots are histograms showing the absolute number of self-reported degree k (including zeros). The lines connect the PMF values for the OLS discrete Pareto fit from Liljeros et al. (2003b) and the Yule with $\rho = 2.54$. Note the extensive heaping and the poor quality of the fits of these models.

here as Newman's interest is in epidemic thresholds based on a particular model for the sexual network.

We should also note that the primary interest is in the number of sexual partners rather than the self-reported number. The former is important for determination of epidemic thresholds, while the later has marginal secondary interest in terms of the pattern of bias and reliability.

Figure 2 represents a histogram of the lifetime data for the Swedish men.

It is clear that there is substantial rounding of the self-reported responses(Morris, 1993). Most of this is to round figures, e.g., 10, 20, 30, ..., 100, ..., with some to secondary figures, e.g., 5, 15, 25, ..., and some to 12 and 24 values. Such so-called "coarsening" or heaping of self-reports is natural for open-ended interview questions. It has been extensively studied (see, e.g, (Roberts and Brewer, 2001) for a review) and is present in a wide range of social science data (Handcock and Morris, 1999). In particular (Wiederman, 1997) conducts an in-depth analysis for the lifetime number of sex partners. Wiederman notes that respondents to sensitive questions

demonstrate a marked preference for "round" numbers, but such unreliability does not seem to manifest consistently as either over-reporting or underreporting (Marquis et al., 1986).

Clearly estimation methods that do not take this into account may produce inaccurate estimates and incorrectly assess the quality of the estimates.

We now describe a procedure to take this into account. One advantage of likelihood methods is their ability to be adapted to account for measurement error processes such as rounding. Recall that the likelihood reflects the probability of observing what we have, given a model for the phenomenon. If a respondent reports a value of 20, we may infer that this is likely to be a rounded value and interpret it as a broader statement of the actual number of partners. For example, we could interpret this as a value from 16 to 24 inclusive. This is a means of capturing the unreliability of the reported value, while still indicating the overall level of the response. Each reported value can be ascribed a separate reliability. For example, a report of 1 may be taken as (close to) reliable, while a value of 100 could be taken as a value from 75 to 125. We note that many refinements of this approach are possible, including data-adaptive reliability schemes. However, here we will use a scheme that treats larger rounded values as less reliable than small values. Methodologically, this can be directly incorporated into the likelihood function. For technical details, see Handcock and Jones (2003a).

It is quite difficult to see make graphical comparisons between the fitted models in Figure 2 because of the well-known skewness of the responses. Handcock and Morris (1999) develop a general framework for comparative distributional analysis based on the concept of a *relative distribution*. The relative distribution provides a graphical display that simplifies exploratory data analysis and a statistically valid basis for the development of hypothesis-driven summary measures The relative distribution is the set of percentile ranks that the observations would have in a reference distribution. In this case it is the percentile ranks of the count of sexual partners within a reference distribution (e.g., Yule with $\rho = 3.54$.) Handcock and Morris (1999) is a book–length treatment of relative distribution theory, and include historical references. The variant for discrete data is considered in Chapter 11.

While we can display the relative distribution of the self-reported values to as estimate of the distribution, it is preferable to take into account the rounded processes. A simple way to do this is to consider ranges of counts (e.g., $1, 2, 3, 4 - 6, 7 - 13, \dots, 75 - 100, > 100$). More sophisticated version of this are possible but it serves for exploratory data analysis.

Figure 3 shows the PMF for the relative distribution of the cumulative lifetime sexual partners using two Yule models as the reference distribution. The smooth line encodes the relative frequency of the counts to the value expected under the reference distribution for each of the above ranges of counts. The value of this ratio is shown on the vertical axis. The top axis shows the the cumulative proportion under the reference, and the bottom the midpoint of the group. The long-dashed line at unity indicates what we would expect if the two distributions were identical. We can see that there are fewer self-reports of 1 and 2 than expected under the Yule model, and that there are many more self-reports than expected in the upper tail (degrees above 3). The over-abundance increases until there are over 15 times as many self reports than expected in the range above 20.

These exploratory tools are helpful to see the deviation between the observed values and proposed models. Liljeros et al. (2003a) also analyze the lifetime data, and criticize our analysis:
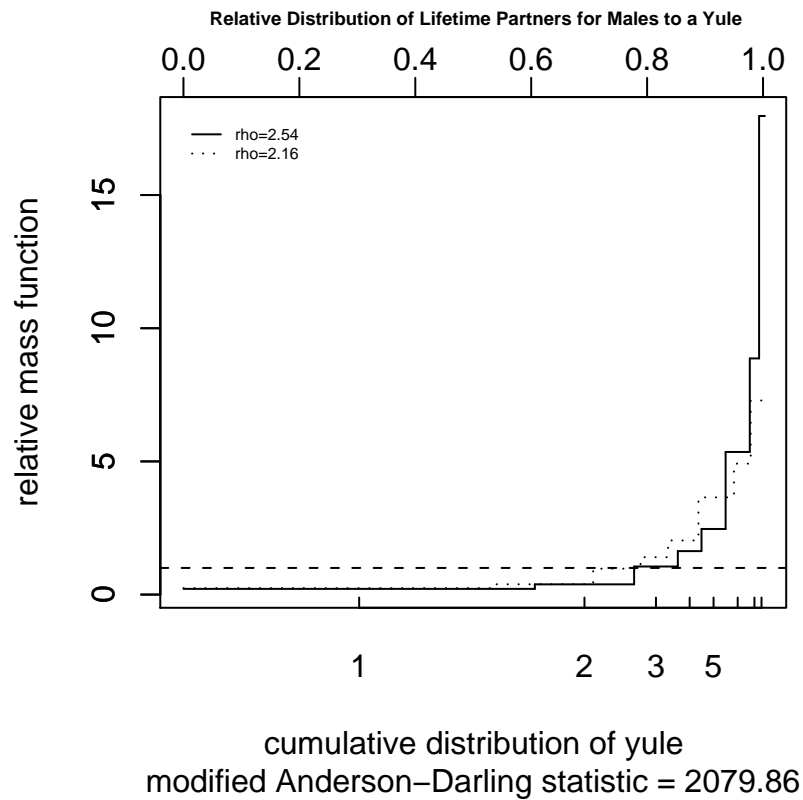
Figure 3: Relative probability mass function for the cumulative lifetime degree distribution to Yule reference distribution. Data are for Swedish men. The plots are PMFs showing the absolute number of self-reported degree k (including zeros). The lines are two power-law fits to the PMF. Note the effect of the heaping and rounding on the upper tail, and the poor fit to the lower tail.

1. "Jones and Handcock also choose to ignore two of the datasets analyzed in Ref. 2, those concerning the number of lifetime partners. Such a practice of limiting one's analysis to a subset of the data (viz., number of partners during a one year period) places unnecessary restrictions on understanding the structure of the network of sexual contacts ..."

   The reason we chose not to report the lifetime partners are given in the cited paper. There are good scientific reasons to be wary of the use of recorded lifetime partners values which we discuss above. In our paper, we chose to consider data from three separate countries rather than just one country. We could ask why Liljeros *et al.* did not consider these, or other countries, in their analysis, but that would be as facile a criticism as theirs.

2. "the analysis in Ref. 2 provides striking support for the hypothesis that the distributions of number of distinct sexual partners decay as a power law."

   The advantage of the graphical methods we use above is that they enable us to see clearly the gross fit of candidate models to the data, especially in the presence of substantial rounding.

   The best fitting Yule distribution for $k_{min} = 20$ has $\rho = 2.54$. As would be expected, the best fitting discrete Pareto is very similar with $\rho = 2.50$. Based on the BIC the Yule is a slightly better fit.

   Figures 2 and 3 display the fit of these models, and a model with slightly lower exponent ($\rho = 2.16$) given in Liljeros et al. (2003b). We can see there is a substantially poor fit to the bulk of the distribution. This visual impression is supported by a Monte-Carlo significance test that rejects the hypothesis of equality.

   Figure 4 focuses in on the upper tail. It graphs the relative mass function for the data above 20 to the Yule distribution conditional on being above 20. As expected, the fit is substantially improved. However, the power-law models struggle to capture the clustered nature of the data especially at 50 and 100. Again, the visual impression is supported by a Monte-Carlo significance test that rejects the hypothesis of equality in both cases.

   These figure show that in attempting to fit the extreme tail of distribution the bulk of the distribution is poorly fit. In addition the rounding and heaping confound the estimation even for this tail.

   It is worthwhile noting that under the best fitting power-law model ($\rho = 2.54$), the tails above 20 contain only 1.23% of the people and 25.3% of the sexual partnerships. Models that accurately represent the other three-quarters of the sexual partnerships would have epidemiological value.

   Similar conclusions hold for the Swedish women. They generally self-report fewer partners. This is a well-known phenomenon (Morris, 1993; Wiederman, 1997). We will report results elsewhere (Jones and Handcock, 2002).

   Finally, repeating the process in Handcock and Jones (2003b) and Jones and Handcock (2003b) we can obtain the best fitting power-law model among those considered in those papers. These are graphically displayed in Figure 5.

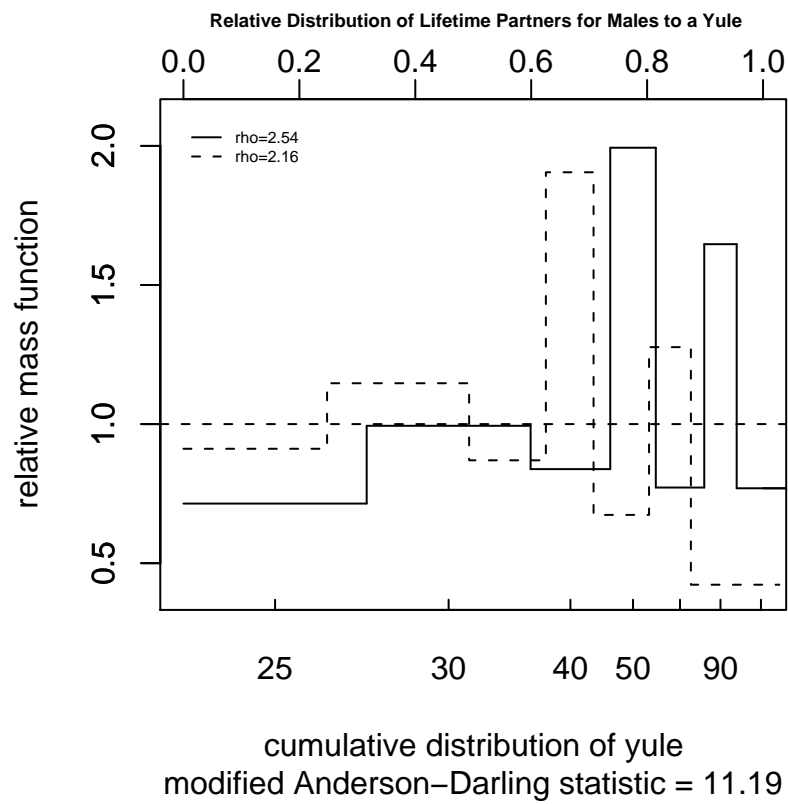**Relative Distribution of Lifetime Partners for Males to a Yule**

Figure 4: Relative probability mass function for the cumulative lifetime degree distribution to Yule reference distribution conditional on degree above 20. Data are for Swedish men.
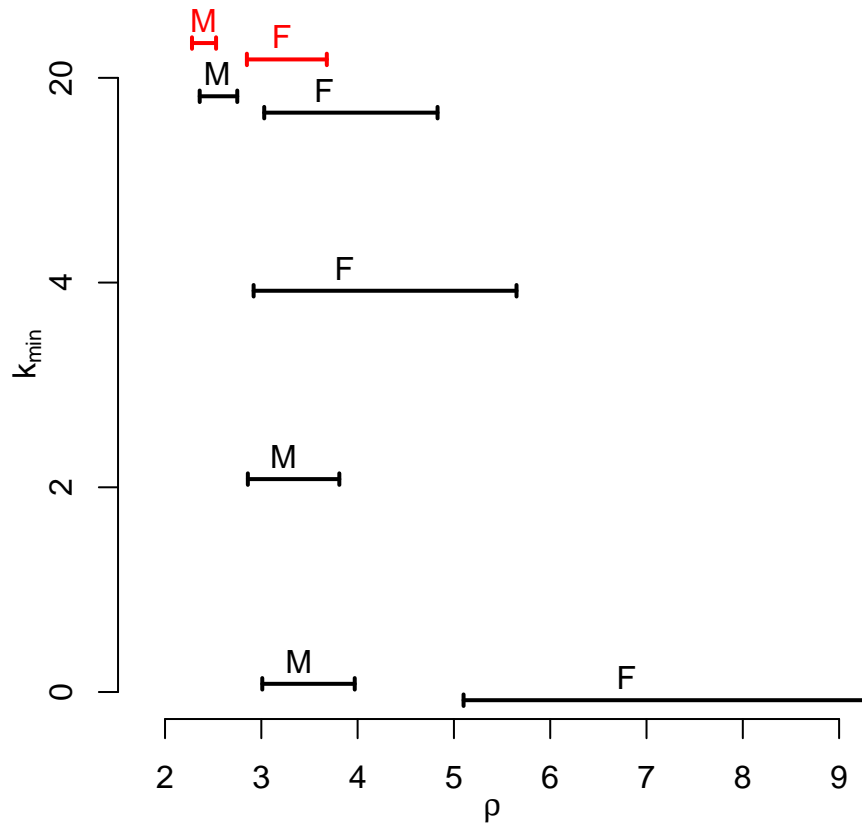
Figure 5: Interval estimates of the scaling parameter, ρ, for the generalized power-law probability mass functions for Swedish males and females. Data show the sensitivity of the 95% bootstrap confidence intervals to the upper tail of the partnership distribution, defined as $k > k_{min}$, where k is the cumulative lifetime number of sex partners. For males the best-fitting model has $k_{min} = 2$, and for females $k_{min} = 4$. By contrast, Liljeros et al. used $k_{min} = 20$ for males and females. For these values, our estimates of parameter uncertainty are three times greater than their reported intervals which we have marked in red.

The best fitting models are Waring for both males and females. The models based on $k_{min} = 20$ provide substantially worse fit based on the BIC.

(a) "However, Jones and Handcock's conclusion is somewhat premature since research convincingly demonstrates that contagious processes differ fundamentally between networks with power law decaying degree distributions and networks with fast decaying degree distributions (Newman, 2002; May and Lloyd, 2001). Specifically, if a detectable epidemic threshold exists, it will certainly be very small as the variance of the distribution of number of sexual partners is much larger than the mean [13-15]."

The two works Liljeros *et al.* cite to support this claim (Newman, 2002; May and Lloyd, 2001) both make it quite clear that the thing that makes "scale-free" networks qualitatively different is infinite variance of the degree distribution. For example, May and Lloyd (p. 2):

> "For the scale-free distributions considered here, CV is infinite because the variance of the connectivity distribution is infinite. In contrast with the case of a homogeneous network, which exhibits threshold behavior at $\rho_0 = 1, R_0$ for the scale-free network (at least for the infinite population case) is infinite for any nonzero transmission probability and so an outbreak can always occur."

As noted above, Newman's paper is full of discussion of epidemic thresholds. He presents analytic formulae for calculating epidemic thresholds. We should note that he never once uses the term "scale-free" in the paper. He too is quite explicit under which circumstances critical behavior of the epidemic will disappear (p. 9):

> "...if the degree distributions are truly power law in form, then there exists an epidemic transition only for a small range of values of the exponent of the power law... if $[\rho] < 3, T_c = 0$ ... As long as [the transmissibility] is positive, we will always be in epidemic regime, and this would clearly be bad news."

Similarly, Dezső and Barabási (2002, 1) note that the key reason scale-free networks are of epidemiological significance is that they are characterized by infinite variance of the degree distribution:

> "*Curing the hubs.* The vanishing epidemic threshold of a virus spreading in a scale-free network is rooted in the infinite variance of the degree distribution."

Non-specialists should be aware that without infinite variance, the relevance of "scale-free" properties for epidemic spread is greatly reduced. In a network characterized by finite variance of the degree distribution, in what way – specifically – do "contagious processes differ fundamentally for scale-free and single-scale networks"? The two works cited in support of this statement focus on infinite variance of the degree distribution.

Our work indicates that the confidence intervals are in fact many times larger than what Newman termed "quite large."

4   Conclusion

Despite the outstanding issues discussed above, there are statements that most people could agree on:

1. Degree distribution of sexual partnerships tend to be right skewed. This will influence the diffusion of STDs.

2. Infinite variance sexual partner degree distributions do not exist in the real world.

3. As the real-world population is finite, mathematical models based on infinite population approximations will need further justification before their conclusions can be applied to determine policy.

4. Epidemic thresholds exist, but some populations may be close to the thresholds.

5. Models which are based purely on information on the degree distribution require substantial further justification before they can be used to determine policy. Nodal attributes, such as age, gender, ethnicity or marital status, are also of fundamental importance for the formation of networks (Morris, 1991). Successful mathematical models will be required to incorporate more information about the population beyond degree.

6. The structure of real-world sexual networks is complex.

It would be productive to agree on common ground and work on the many remaining issues. We believe that it is sufficient to foster collaboration, and we intend to focus our energies on productively working on those issues.

In sum, the results presented in our work and others indicate that a general behavioral model for the formation of human sexual contact networks is still lacking. Preferential attachment is one model that has been previously suggested, and the assumptions of preferential attachment have been incorporated into our work. However, this mechanism does not perform especially well when confronted with alternative models. We suggest that incorporating actor heterogeneity and the timing and sequence of partnerships is essential for future network models in epidemiology (Handcock, 2002; Morris, 2002). Furthermore, if we are to move beyond *ad hoc* curve fits of network degree distributions and make real progress in understanding the stochastic mechanisms which generate empirical networks, two points are essential: (1) we must recognize that there is much more to sexual networks than degree distributions, and (2) collaboration between network modelers, epidemiologists, and social scientists is essential.

References

Albert, R. and A. L. Barabási (2000). Topology of evolving networks: Local events and universality. *Physical Review Letters 85*(24), 5234–5237.

Amaral, L. A. N., A. Scala, M. Barthelemy, and H. E. Stanley (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America 97*(21), 11149–11152.

Anderson, R. M. and R. M. May (1991). *Infectious diseases of humans: Dynamics and control.* Oxford: Oxford University Press.

Berger, J. (1986). *Statistical decision theory and Bayesian analysis.* New York: Springer-Verlag.

Brauer, F. and P. van den Driessche (2002). Some directions in mathematical epidemiology. In C. Castillo-Chavez, S. Blower, D. van den Driessche, P. andKischner, and A. Yakubu (Eds.), *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, Volume 126, pp. 302–322. New York: Springer Verlag.

Dezső, Z. and A. L. Barabási (2002). Halting viruses in scale-free networks. *Physical Review E 65*, art. no. 055103.

Estroup, J. (1916). *Les gammes sténographiques.* Paris: Institut Sténographique.

Farmer, P. (1999). *Infections and inequalities.* Berkeley: University of California Press.

Farmer, P. (2003). *Pathologies of power: Health, human rights, and the new war on the poor.* Berkeley: University of California Press.

Ford, N. and S. Koetsawang (1999). A pragmatic intervention to promote condom use by female sex workers in Thailand. *Bulletin of the World Health Organization 77*(11), 888–94.

Handcock, M. and J. Jones (2003a). Confidence bounds on two-sex epidemic thresholds. *Proceedings of the Royal Statistical Society, A submitted.*

Handcock, M. S. (2002). Statistical models for social networks: Inference and degeneracy. In R. Breiger, K. Carley, and P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis*, Volume 126, pp. 302–322. Washington, DC.: Committee on Human Factors, Board on Behavioral, Cognitive, and Sensory Sciences, National Academy Press.

Handcock, M. S. and J. Jones (2003b). Likelihood-based inference for stochastic models of sexual network formation. *Theoretical Population Biology submitted.*

Handcock, M. S. and M. Morris (1999). *Relative Distribution Methods in the Social Sciences.* New York: Springer-Verlag.

Johnson, N., S. Kotz, and A. Kemp (1992). *Univariate discrete distributions* (2nd ed.). Wiley series in probability and mathematical statistics. New York: Wiley.

Jones, J. and M. S. Handcock (2002). Assessing sexual network degree with the relative distribution.

Jones, J. and M. S. Handcock (2003a). An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society of London, B 270*, 1123–1128.

Jones, J. and M. S. Handcock (2003b). Sexual contacts and epidemic thresholds. *Nature in press*.

Katz, B. P. (1992). Estimating transmission probabilities for chlamydial infection. *Statistics in Medicine 11*(5), 565–577.

Kim, J. Y., J. Millen, J. Gershman, and A. Irwin (Eds.) (2000). *Dying for Growth: Global Inequality and the Health of the Poor*. Monroe, ME: Common Courage Press.

Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg (2001). The web of human sexual contacts. *Nature 411*(6840), 907–908.

Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg (2003). Authors' reply. *Nature in press*.

Liljeros, F., C. R. Edling, H. E. Stanley, Y. Åberg, and L. A. N. Amaral (2003b). Distributions of number of sexual partnerships have power law decaying tails and finite variance. *ArXiv Preprint cond-mat/0305528*.

Marquis, K. H., M. S. Marquis, and J. M. Polich (1986). Response bias and reliability in sensitive topic surveys. *Journal of the American Statistical Association 81*(394), 381–389.

May, R. M. and A. L. Lloyd (2001). Infection dynamics on scale-free networks. *Physical Review E 64*(6), 066112.

Morris, M. (1991). A log-linear modeling framework for selective mixing. *Mathematical Biosciences 107*(2), 349–377.

Morris, M. (1993). Telling tails explain the discrepancy in sexual partner reports. *Nature 365*(6445), 437–440.

Morris, M. (2001). Concurrent partnerships and syphilis persistence: New thoughts on an old puzzle. *Sexually Transmitted Diseases 28*(9), 504–507.

Morris, M. (2002). Local rules and global properties: Modeling the emergence of network structure. In R. Breiger, K. Carley, and P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis*, Volume 126, pp. 302–322. Washington, DC.: Committee on Human Factors, Board on Behavioral, Cognitive, and Sensory Sciences, National Academy Press.

Morris, M., P. C., W. M. J., and H. M. S. (1997). Bridge populations in the spread of HIV/AIDS in thailand. *AIDS 10*(11), 1265–1271.

Neaigus, A. (1999). The network approach and interventions to prevent HIV among injection drug users. *Public Health 113*, 140–150.

Nelson, K. E., S. Eiumtrakul, D. D. Celentano, C. Beyrer, N. Galai, S. Kawichai, and C. Khamboonruang (2002). HIV infection in young men in northern Thailand, 1991-1998: Increasing role of injection drug use. *Journal of Acquired Immune Deficiency Syndromes 29*(1), 62–8.

Newman, M. E. J. (2002). Spread of epidemic disease on networks. *Physical Review E 66*(1), art. no.–016128.

Roberts, J. M. and D. D. Brewer (2001). Measures and tests of heaping in discrete quantitative distributions. *Journal of Applied Statistics 28*(7), 887–896.

Tanur, J. M., R. S. Pieters, and F. Mosteller (1989). *Statistics: A Guide to the Unknown*. New York: Brooks-Cole.

Tuljapurkar, S., N. Li, and C. Boe (2000). A universal pattern of mortality decline in the g7 countries. *Nature 405*(6788), 789–792.

Wiederman, M. W. (1997). The truth must be in here somewhere: Examining the gender discrepancy in self-reported lifetime number of sex partners. *Journal of Sex Research 34*(4), 375–386.

Zipf, G. (1949). *Human behavior and the principle of least effort*. New York: Addison-Wesley.