Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks[1]

Peter S. Bearman
Columbia University

James Moody
Ohio State University

Katherine Stovel
University of Washington

This article describes the structure of the adolescent romantic and sexual network in a population of over 800 adolescents residing in a midsized town in the midwestern United States. Precise images and measures of network structure are derived from reports of relationships that occurred over a period of 18 months between 1993 and 1995. The study offers a comparison of the structural characteristics of the observed network to simulated networks conditioned on the distribution of ties; the observed structure reveals networks characterized by longer contact chains and fewer cycles than expected. This article identifies the micromechanisms that generate networks with structural features similar to the observed network. Implications for disease transmission dynamics and social policy are explored.

INTRODUCTION

This article describes the structure of adolescent romantic and sexual networks in an American high school, accounts for the emergence of this

[1] Data for this article are drawn from the National Longitudinal Study of Adolescent Health (Add Health), a program project designed by J. Richard Udry and Peter Bearman, and funded by a grant from the National Institute of Child Health and Human Development (HD31921). The authors thank Douglas White, Martina Morris, Mark Handcock, J. Richard Udry, and the *AJS* reviewers for helpful comments on previous drafts of this article. Direct correspondence to Peter Bearman, Institute for Social and Economic Research and Policy, 814 SIPA Building, Columbia University, New York, New York 10027. E-mail: psb17@columbia.edu

© 2004 by The University of Chicago. All rights reserved.
0002-9602/2004/11001-0002$10.00

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

structure, and links the observed structure to the diffusion dynamics of disease. Our goal is to show how local preferences governing partner choice shape the macrostructures in which individuals are embedded and hence affect both the potential for disease diffusion and the determinants of individual risk.[2] Because the structure of sexual networks is critical for understanding the diffusion of sexually transmitted diseases (STDs), it is surprising that epidemiologists have only a limited idea of what such networks look like. The insight we do have is generally restricted to that provided by a set of ego-centered network surveys (Morris and Kretzschmar 1995, 1997; Laumann et al. 1994; Laumann and Youm 1999) and snowball samples of populations of highest risk to HIV acquisition, such as male homosexuals (Klovdahl 1985) and IV drug users (Rothenberg, Potterat, and Woodhouse 1996; Rothenberg et al. 1997; Friedman et al. 1997). While they may reveal much about the characteristics of the local networks in which individuals are embedded, ego-centered and snowball samples provide limited information on the global network properties that determine disease spread.

In this article, we describe extensive partnership patterns and network

[2] As background, each year in the United States over 12 million individuals discover that they carry a sexually transmitted disease (STD). The two leading STDs, herpes and human papillary virus (HSVT2 and HPV, respectively), are chronic and, although subject to palliative treatment, not curable. Adolescent STD acquisition rates outpace those of all other groups, with no change in sight. Roughly 5% of all sexually active adolescents have acquired chlamydia or gonorrhea (Aral et al. 1999). Among sexually active black adolescents, 25% are likely to be infected with herpes (CDC 2000), and probably 40%–50% of all sexually active females have had a previous HPV infection, now known to account for most cases of adult cervical cancer (Holmes et al. 1999). The literature identifies three reasons for these gloomy facts. First, one-half of all adolescents over 15 years old report being sexually active, and a significant proportion of these adolescents are inconsistent in their use of condoms, therefore heightening risk of STD acquisition and transmission (Bearman and Brückner 1999). Furthermore, many adolescents who have not had intercourse are sexually active in a substantively meaningful (if technically ambiguous) way, and most do not use condoms during noncoital sex. Specifically, of adolescents who report that they are virgins (i.e., have not had sexual intercourse) roughly one-third have had genital contact with a partner resulting in fluid exchange in the past year. Thus virginal status does not mean that adolescents are not engaging in behaviors that are free of risk for STD transmission. Second, the majority of adolescents with an STD have no idea that they are infected (Holmes et al. 1999); consequently, they may fail to protect their partners even if they would prefer to do so. And third, relative to adults, adolescents tend to form romantic partnerships of short duration, on average only 15 months, but with a strong skew toward relationships of extremely short duration (less than four months; Laumann et al. 1994). Most sex in adolescent relationships occurs, if it is to occur, within the first two months (Bearman, Hillmann, and Brückner 2001). This combination of short duration partnerships, inconsistent safe-sex practices, and incorrect assessment of STD status provides a partial account for the diffusion of STDs among the adolescent population. As fundamental is the role that sexual contact structures play in STD transmission dynamics.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

structure for one population of interacting adolescents in a midsized American town, thereby providing detailed images of, and measurement for, key structural characteristics of a largely complete romantic and sexual network through which STDs may diffuse. As background, we begin by describing some models of sexual networks that are implicit in the existing literature on STDs. We then report the structure of the network generated by the romantic and sexual partnership nominations provided by most of the adolescents in the study community. We consider both cross-sectional and temporal views of this network, and we discuss the extent to which the cross-sectional view obscures the potential for disease diffusion. We then turn to how such a structure could emerge. Because it is theoretically possible that homophily in partner selection—the tendency for individuals with similar attributes, characteristics, or practices to form partnerships—could generate the network structure we observe, we explore the determinants of partnership choice and show that the observed structure is not *solely* a by-product of preferences for particular attributes. We subsequently propose a parsimonious micromodel that, given the determinants of partnership choice, accounts for the structure we observe. Implications for public policy are considered in the conclusion.

Below, we show that (1) current models of disease diffusion rest on sexual network structures that differ in fundamental ways from what we observe, (2) preferences governing partner choice combined with a simple normative proscription against cycles of length 4 (Don't date your old partner's current partner's old partner) induce the structure we observe, (3) partnership preference models that ignore the proscription against completing cycles of length 4 induce incorrect structural representations, and (4) consequently, current intervention efforts that assume the existence of cores may be poorly conceived.

MODELS OF DISEASE DIFFUSION

The fundamental quantity in models of disease diffusion is the basic reproductive rate R_o.[3] When $R_o > 1$, a self-sustaining epidemic occurs; when $R_o < 1$, the disease dies out. In models of disease diffusion, the reproductive rate is a function of three parameters: the infectivity of the microbe given contact between an infected and a susceptible (β), average duration of infectiousness (D), and the structure of disease-relevant contact within a population (C). The critical sociological parameter is C, the network structure that governs contact.

[3] R_o is defined as the number of new infections produced by an infected individual over the duration of infectivity (Anderson and May 1991).

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

The simplest epidemiological models assume *random mixing* among all members of the population. Under random mixing, the number of new infections at time t is easily calculated as the number of susceptibles times the number of infecteds times the proportion of contacts between susceptibles and infecteds that result in infection. The result of a random mixing model is the classical S-shaped diffusion curve, where one observes a slow start, followed by exponential growth, and then a decline, either from recovery or death (Sattenspiel 1990).[4]

One can think of random mixing as the statement "people choose partners independent of their characteristics." For many diseases, random mixing captures the essential aspects of the diffusion process. The sneeze of a flu-ridden person on a transatlantic plane sends viral and bacterial material through the air, potentially infecting all of the passengers, though those sitting next to the sick person are at greatest risk. Although we may feel otherwise in our less gracious moments, we know that the airlines did not select *us* to sit next to a sneezer and that he or she did not sneeze on us because of our characteristics. For STDs, however, pure random mixing provides a poor approximation of the underlying contact structure.[5]

As sociologists have long noted, partner-selection processes count. Thus models that explicitly consider bias in partner choice may more closely reflect the social and behavioral processes that give rise to disease-relevant contact structures. For example, the obvious bias relevant for diseases spread via heterosexual contact is toward partners of the opposite sex. Among two-sex models of disease diffusion, the best-known class of partner-bias models are preferred-mixing models that assume disproportionately high levels of contact between individuals who share some attribute (Koopman et al. 1989; Sattenspiel 1990; Jacquez et al. 1988; Hethcote 2000).[6] Based on the homophily principle, these models recognize that, given opposite-sex partnerships, persons often prefer contact with those who are similar to themselves with respect to race, religiosity, sexual

[4] The S-shaped curve is not specific to random mixing. As noted by a reviewer, the S shape may result from many different contact structures.

[5] This is not to suggest that such models have no utility. For example, as one reviewer notes, mixing models in which groups are based on the number of sexual partners show that the probability of having sex with an infected person exceeds the prevalence of infection in the population. However this would not be a *pure* random-mixing model.

[6] Preferred-mixing models operate on persons classified by attribute rather than by structural position. The models we develop subsequently identify position as the critical element, such that attributes of persons are substitutable across positions. But, in an abstract sense, they are also preferred-mixing models.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

preference, activity level, and so on.[7] In such models, leftover contacts occur between people of different groups proportional to the level of sexual activity of these groups. Depending on the values of specific mixing parameters, these preferred-mixing models predict different levels and patterns of disease spread.

Though systematic differences in connectivity patterns have been shown to have striking implications for disease transmission (Morris 1997; Newman 2002; Dezso and Barabasi 2002; Moody et al. 2003), preferred-mixing models do not consider sexual network structure in a direct way. Yet representing the models as networks is a useful way to reveal their assumptions about contact patterns.[8] Three stylized images of sexual networks can be derived from the literature on the diffusion of STDs.[9] The first, and most influential, is that of a *core*. According to standard models, a core is a group of high activity-level actors (e.g., those with multiple partners or who are frequent drug users) who interact frequently and pass infection to one another (often causing reinfection for treatable STDs), and diffuse infection out to a less densely connected population (Phillips, Potterat, and Rothenberg 1980; Hethcote and Yorke 1984*a*, 1984*b*; St. John and Curran 1978). Under the general diffusion model, cores are predicted to sustain endemic pockets of disease, since the pattern of intense interaction among members of the core pushes R_0 in the core above 1.

We represent the network structure implied by a core model in figure 1a; here, circles represent individuals and lines represent (disease-relevant) relationships. High activity actors (core members) are indicated by black circles, and the core is circled. Here we do not differentiate by sex: core membership is determined by activity level, and the core is assumed to contain both males and females. Translating a core-based preferred-

[7] Preferred-mixing models need not assume homogeneous mixing within groups. Variants include assortative-matching models that incorporate out-group preferences, e.g., age-skewed models that match older males with younger females, or role separation models. The implications of skewed age matching for HIV diffusion are explored in Morris (1993).

[8] An important development has been the focus on strongly skewed positive degree distribution in scale-free networks for diffusion dynamics (Newman 2002; Dezso and Barabasi 2002; Newman et al. 2001; Barabasi and Albert 1999; Watts 2003). As noted subsequently, the structure we observe is not a scale-free network with a power law distribution of degree. Moody et al. (2003) show that a large densely connected core can emerge in populations with low degree, demonstrating that the structural conditions for large epidemics are possible even in populations without skewed partnership distributions.

[9] Each of these models has been refined in the literature to incorporate both theoretical and empirical advances in our understanding of the process by which sexual partnerships are formed. Thus, e.g., there are variants of these models that emphasize role separation in male homosexual populations, temporal overlap of partnerships (concurrency), and multiple or complex preference structures in heterogeneous populations.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

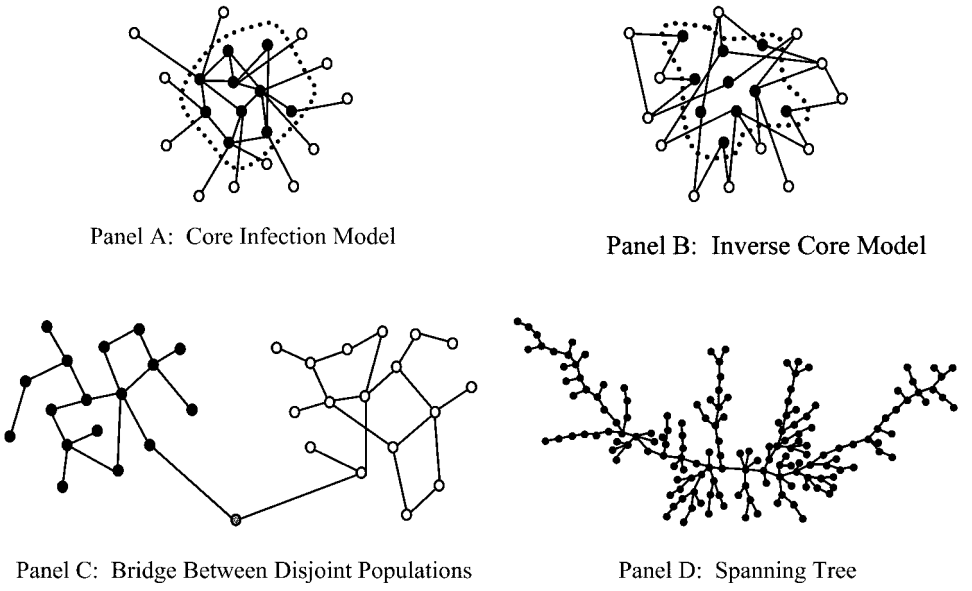

FIG. 1.—The network structure of four models of infection

mixing model into a network structure highlights specific measurable properties of the resulting graph. In a core, it is likely that an individual's past partner is tied through multiple chains to his or her current or future partner. Thus if cores exist in a population, cyclicity will be extremely high in the network, and the length of chains connecting pairs of individuals in the population (geodesics) will be low.

While core-based models have been used to account for the diffusion of bacterial STDs like gonorrhea (Hethcote and Yorke 1984*a*, 1984*b*; Hethcote and van den Driessche 2000; Hethcote and Van Ark 1987; Aral et al. 1999; St. John and Curran 1978), core models offer a poor description of sexual contact patterns in many contexts. For example, when a key mode of transmission is male long-distance truck drivers having sex with female commercial sex workers (CSWs), members of the groups that constitute possible infection reservoirs (like CSWs) are structurally disconnected from one another and do not transmit infection *directly* to one another. Capturing such dynamics—which may be more characteristic of two-sex diffusion processes—requires more complex switching models, often called *inverse core* models (Garnett et al. 1996). In an *inverse core*, a central group of infected persons pumps disease out to others but does not pass infection directly among themselves. For instance, prostitutes might be infected by previously infected johns and then pass infection on to other johns.

We represent the network associated with an inverse core in figure 1b. Here we distinguish actors according to their *role* in the diffusion process (commercial sex workers are black, sex customers are white; the inverse core is circled). The key difference between a core and an inverse core stems from the social organization of sexual relations, since johns are more likely than other potential carriers to spread infection to individuals not in the graph (specifically, their regular sex partners). Structurally, however, the two networks are quite similar; though cores may be smaller and denser, both structures are associated with high cyclicity and low path distance between individuals. Since viruses are hardly attentive to the social details that occupy us, the two structures hold similar potential for disease diffusion.

A third model in the epidemiological literature describes disease diffusion dynamics as driven by *bridging processes* (Aral 2000; Gorbach et al. 2000; Morris et al. 1996). These models posit two populations of persons engaged in different behaviors (i.e., a high-risk and a low-risk population) linked by a few individuals who bridge the boundary between each world (e.g., an IV drug user who shares needles with his drug partners and who has sex with non-IV-drug users). A network consistent with this model is shown in figure 1c, where the black circles denote actors who engage in

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

high-risk behavior and the white circles denote actors who engage in lower-risk behavior. Bridges are those who link the two worlds.[10]

Each of these graphs represents the network foundation of a preferred-mixing model. In the core and inverse core models, mixing results from in-group preference with respect to risk status (or some other attribute, like IV drug use, associated with risk status). Even in the bridging model, parameters for cross-group contacts are estimated from individual-level data, and random mixing within groups is presumed. In all cases, disease potential is contingent upon the extent to which (1) there is at least one local pocket of densely interconnected persons (connected either via direct connections as in the core or via short and redundant cycles as in the inverse core) and (2) the pocket of high density is connected to the remainder of the population through bridging chains that reach into the periphery of the network. However, these models are useful only to the extent that any empirically observed network structure matches that implied by the in- and out-group contact parameterization. Obviously, if the network one actually observes bears little relationship to the structure implied by these models, we must radically reconsider their usefulness.[11] For our data, this is the situation.

Specifically, in the context we study we observe a network structure that has the appearance of a *spanning tree*; that is, a long chain of interconnections that stretches across a population, like rural phone wires running from a long trunk line to individual houses (Hage and Harary 1996, 1983). The global structure of a chainlike spanning tree is characterized by a graph with few cycles, low redundancy, and consequently very sparse overall density.[12] The shortest distance between any two randomly selected individuals (geodesic) is significantly higher than that observed in either the core or inverse core structures.[13] A typical spanning tree structure is represented as figure 1d.

Random-mixing dynamics and positive preferences for partners do not

[10] This graph expands on the triads linking the core to the periphery in the core and inverse core networks, though this image draws attention to the fact that the bridging triad is embedded in a macrostructure different from either of these other two models.

[11] Recent work on the structure of large networks—e.g., those linking nodes on the World Wide Web—reveals starlike nodes with very high degree. Such a structure is not replicated in the data we observe.

[12] Here and throughout we refer to chainlike structures that are not dominated by cycles or small numbers of highly central nodes as spanning trees, though technically a spanning tree is any connected noncycle graph.

[13] A super star graph is technically a spanning tree, but is associated with short geodesics (Barabasi and Albert 1999).

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

produce spanning tree structures.[14] Rather, this network structure appears when formal or informal rules *preclude* the enactment of specific relations. In the language of kinship structures, spanning trees are the product of negative proscriptions: sets of rules about whom one *cannot* be in a relationship with. Consequently, they are most frequently observed in large and complex generalized exchange systems, as in the exchange of valuables in the Kula ring (Hage and Harary 1996; Schweizer and White 1998).

As noted above, the extant models of sexually transmitted disease diffusion implicitly assume network structures that correspond to one of the first three images in figure 1. Yet we have essentially no complete population data from which to conclude that any of these models are empirically appropriate. Fundamental at this point is the need to learn more about actual networks and the structural characteristics that are relevant for disease diffusion. In this article, we describe these characteristics in an observed romantic and sexual network in a population of adolescents. The network structure we find closely approximates a spanning tree. Since such structures are the result of rules restricting partnership choice, we focus on identifying a parsimonious rule that could produce the structure we observe empirically. Conditional on simple homophily preferences in partnership choice, the structural properties of networks simulated according to this rule closely correspond to what we observe.

CONTEXT AND DATA

Data for this article are drawn from the wave 1 component of the National Longitudinal Study of Adolescent Health (hereafter, Add Health), a longitudinal study of adolescents in grades 7–12. In 1994, in-school questionnaires were administered to approximately 90,000 students in 140 schools. Almost a year later, a nationally representative sample of over 20,000 of these students completed extensive interviews in their homes. In 14 saturated field settings composed of two large (N = 1,000; N = 1,800) and 12 small ($N < 300$) schools, Add Health attempted home interviews with *all* enrolled students. The two large schools were selected with the intent of capturing typical high school experiences in urban and less urban communities. The adolescent in-home interview was conducted using audio-CASI technology for all sensitive health status and health risk behavior questions. Adolescents listened to the questions through earphones and directly entered their responses into a computer, thereby

[14] One reviewer suggests that if sex "has an element of contagion in which only sexually experienced actors recruit new participants," the resulting graph would be a spanning tree. See app. C for our response to this suggestion.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

eliminating interviewer or parental effects on their responses (Turner et al. 1998). Adolescents were asked to identify their sexual and romantic partners from a roster of other students attending their school. Consequently, in the saturated field settings, we have almost complete sexual and romantic network data.

Context: "Jefferson High"

In this article we report data from the 832 respondents who attended a school we identify as "Jefferson High School," one of the two large high schools where Add Health attempted in-home interviews with all students. Jefferson High is an almost all-white high school of roughly 1,000 students located in a midsized midwestern town. Jefferson is the only public high school in the town. The town, "Jefferson City," is over an hour's drive from the nearest large city. While densely settled, Jefferson City is surrounded by beautiful countryside and is home to many agricultural enterprises. At one time the town served as a resort for city dwellers, drawing an annual influx of summer visitors, though this is no longer the case, and many of the old resort properties show signs of decay. At the time of the fieldwork, students were reacting to the deaths of two girls killed in an automobile accident. Despite this, fieldwork proceeded exceptionally well. Adolescents frequently approached interviewers wearing yellow *Add Health* buttons and asked when they would be invited to participate in the study.[15] In all, 90% of the students on the school roster participated in the in-school survey, and over the course of the interview period, 83% of all students in the school completed in-home interviews.

Jefferson is a close-knit, insular, predominantly working-class community, which offers few activities for young people. In describing the events of the past year, many students report that there is absolutely nothing to do in Jefferson. For fun, students like to drive to the outskirts of town and get drunk. For our purposes, the relative isolation of the community is an important factor, significant for the patterns of romantic partnership and sexual partnership choices we observe. The context provides a good setting in which to look for the networks suggested by preferred-mixing models, for if redundant structures (and therefore, cores) exist, they are most likely to appear in island populations not permeated by the currents of larger, more cosmopolitan settings.

Table 1 describes the tenth-to twelfth-grade students at Jefferson High across a broad spectrum of characteristics.[16] It also contains comparisons

[15] Adolescents were given $20 in appreciation for completing the interview. Just before Mothers' Day and the prom, many adolescents were eager to be interviewed.

[16] Additional information about these measures is provided in app. A.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

TABLE 1
School-Level Comparisons: Jefferson High and Other High Schools in the Add Health Sample

		High School			
Variable	Jefferson High	All	Mainly White	600–1,000 Students	White, 600–1,000 Students
Family SES	5.59	5.78	5.95	5.73	5.86
		(1.16)	(.91)	(1.20)	(.86)
Proportion in poverty	.13	.16	.09**	.16	.13
		(.14)	(.09)	(.15)	(.13)
Log(family income)	3.61	3.56	3.65	3.55	3.54
		(.42)	(.29)	(.47)	(.35)
GPA	2.49	2.83***	2.93***	2.78***	2.82**
		(.24)	(.18)	(.19)	(.10)
Expect college graduation	3.77	4.09***	4.18	4.09***	4.09*
		(.48)	(.36)	(.30)	(.25)
School attachment	3.27	3.67***	3.68***	3.59***	3.49*
		(.26)	(.29)	(.23)	(.17)
Trouble in school	1.20	1.01***	1.02***	1.04***	1.05*
		(.13)	(.11)	(.12)	(.13)
Drunk	1.14	.86***	.94**	.93***	1.02*
		(.30)	(.30)	(.27)	(.11)
Delinquency	.29	.26**	.25*	.27	.25
		(.07)	(.06)	(.05)	(.05)
Hours watching TV	10.29	13.7***	11.81	13.87**	12.03
		(3.52)	(2.74)	(4.27)	(2.39)
Religiosity (praying)	2.71	2.09***	2.22***	2.18***	2.32
		(.42)	(.49)	(.38)	(.37)
In-degree	5.32	4.39***	4.85*	4.61**	5.00

		(1.02)	(.77)	(.92)	(.41)
Self-esteem	2.99	3.18***	3.17***	3.18***	3.14**
		(.11)	(.09)	(.12)	(.08)
Sexually active	.59	.53	.51	.57	.55
		(.16)	(.16)	(.14)	(.10)
Autonomy	.86	.80***	.82***	.82***	.84*
		(.05)	(.04)	(.05)	(.03)
Expects to get AIDS	1.52	1.53	1.53	1.58	1.58
		(.15)	(.11)	(.10)	(.09)
Marry by 25	3.06	3.18***	3.32***	3.13**	3.21**
		(.25)	(.21)	(.17)	(.13)
Attractiveness	3.45	3.59***	3.59**	3.59	3.57
		(.21)	(.18)	(.22)	(.22)
AH_PVT	105.32	101.05***	105.29	102.06	105.11
		(8.96)	(3.06)	(6.67)	(3.62)
Two biological parents	.46	.53**	.58***	.54	.58*
		(.13)	(.11)	(.13)	(.08)
Smokes regularly	.36	.26***	.32*	.28***	.32*
		(.11)	(.09)	(.08)	(.03)
School suspension	.40	.27***	.22***	.30**	.24**
		(.14)	(.12)	(.14)	(.11)
Tattoo	.10	.06***	.05***	.05***	.06*
		(.04)	(.03)	(.03)	(.03)
Number	1	75	28	23	9

NOTE.—Nos. in parentheses are SDs (Students in tenth, eleventh, and twelfth grades only).
* $P < .05$, whether the median of the sample distribution equals the Jefferson value, based on a sign test.
** $P < .01$.
*** $P < .001$, whether the median of the sample distribution equals the Jefferson value, based on a sign test.

with all other high schools in the sample (col. 2); all disproportionately white schools (over 75% white; col. 3); high schools of comparable size (col. 4); and finally, the small set of other disproportionately white high schools of similar size in Add Health (col. 5).

Sign tests reveal that, in general, Jefferson High is similar to other U.S. schools across most of the comparison variables.[17] However, Jefferson students earn lower grades, are suspended more often, feel less attached to school, and come from poorer families than those at comparable schools. They are more likely than students in other high schools to have trouble paying attention, and they have lower self-esteem, pray more, have fewer expectations about college, and are more likely to have a permanent tattoo. Compared to other students in large, disproportionately white schools, adolescents at Jefferson High are more likely to drink until they are drunk. In schools of comparable race and size, on average 30% of tenth-to-twelfth grade students smoke cigarettes regularly, whereas in Jefferson that figure is 36%. Drug use is moderate, comparable to national norms. More than half of all students report having had sexual intercourse, a rate comparable to the national average and only slightly higher than observed for schools similar with respect to race and size. Jefferson is not Middletown, but it looks an awful lot like it.

Romantic Partnerships and Sexual Partnerships

During the in-home interview, adolescents were asked if they were in or had been involved in a *special romantic relationship* at some point during the past 18 months. Adolescents in such relationships were asked to describe their three most recent relationships, including any current relationships.[18] In addition, adolescents were asked to identify up to three

[17] Reviewers suggested a nonparametric test for these comparisons. Following Conover (1980), we use a sign test to compare the median value of each of these characteristics at Jefferson to each of the four subsets of other schools. The assessment is quite sensitive, especially with respect to col. 5, where the *N* of comparison schools is small. The differences between Jefferson and other schools in the sample can be qualitatively described simply: Jefferson is an all-white school that is largely working class; most all-white schools in the country are composed of upper-middle-class adolescents who reside in segregated suburbs. Consequently, social class predominantly drives differences in behavior, academic orientation, and achievement.

[18] Adolescents who did not identify that they had a special relationship were asked if in any relationship over the past 18 months they had "held hands, kissed, or told someone that they liked or loved them." If an adolescent was in such a relationship, then they were asked to identify their partner and describe their relationship. Both self-identified and behavior-induced "partnerships" could, but did not necessarily, involve sexual intercourse. Of adolescents who reported being a virgin (i.e., had not had sexual intercourse) one-third had had genital contact with a partner resulting in fluid exchange in the past year (Schuster, Bell, and Kanouse 1996). Thus not having inter-

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

individuals with whom they had a *nonromantic sexual relationship* in the past 18 months. A nonromantic sexual relationship was defined as a relationship involving sexual intercourse that the respondent did not identify as special and in which the partners did not kiss, hold hands, or say that they liked each other. A large number of sexual, nonromantic relationships were reported. For the vast majority of reported partnerships, start and end dates for all romantic and nonromantic sexual partnerships were collected. Slightly less than one-quarter of all Jefferson students reported no romantic or nonromantic sexual relationship during the preceding 18 months.

After collecting detailed information about partnerships, respondents were asked if their partner attended their school (or the middle school that fed students into the high school). If their partner attended either school, respondents were asked to identify their partner from a roster by a unique ID. Through this process we collected data on 477 partnerships between respondents at Jefferson High and one of the two sampled schools in Jefferson.[19] We use these partnerships to generate a snapshot of the network of romantic and sexual relations among adolescents attending high school in Jefferson—the first such image that does not rely solely on egocentric reports from a small fraction of the relevant population.

OBSERVED ROMANTIC AND SEXUAL NETWORKS AT JEFFERSON HIGH

Figure 2 maps the actual network structure that links the 573 students involved in a romantic or sexual relationship with another student at Jefferson High.[20] Circles denote individual students; romantic or sexual relations are flows between nodes. Time is suppressed in this representation.

---

course does not mean refraining from behaviors that are risky for HIV or STD transmission (although the noncoital fluid-exchange behaviors they do engage in carry less risk for both partners than intercourse).

[19] Nominations to students account for 51.2% of all romantic nominations and 39.4% of all nonromantic sexual partnership nominations. These partnerships involve roughly 75% of all students who reported having a romantic relationship. In the other large saturated field setting in the Add Health sample, which is located in an ethnically heterogeneous metropolitan area, only 11% of all partnership nominations were directed toward other students. Thus in Jefferson, the school community provides the key focal context (Feld 1981) for adolescent social and sexual relations. Given the relative isolation of the community, this orientation is expected.

[20] In fig. 2, and in all discussions presented here, all romantic and sexual relationship nominations linking students are included, whether or not the nomination from i to j was reciprocated with a nomination from j to i.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

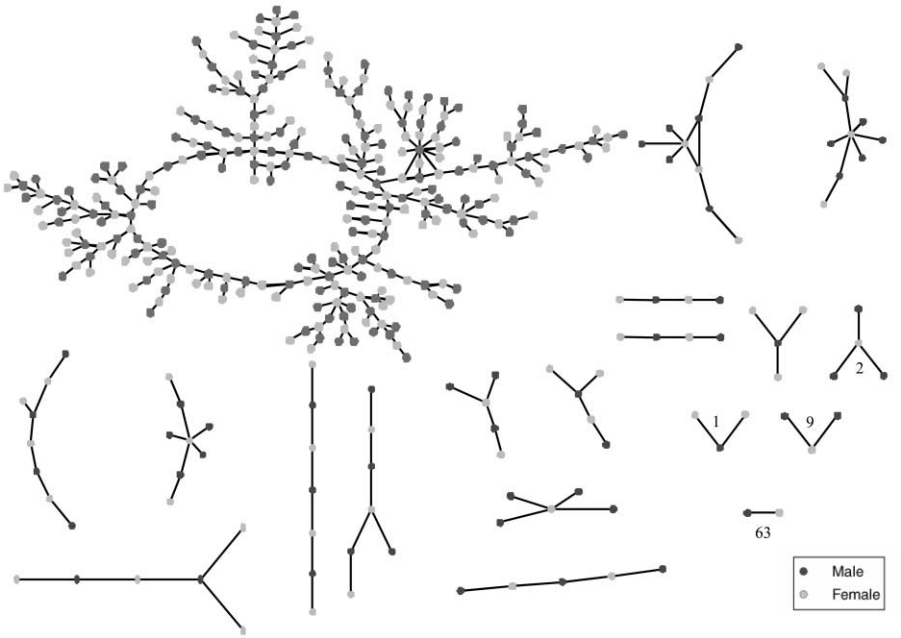

FIG. 2.—The direct relationship structure at Jefferson High

We begin by examining the distribution of components in this network.[21] A component is a subgraph of a network in which all nodes are reachable from other nodes in the subgraph (Wasserman and Faust 1994). Components are significant for disease transmission, since individuals who are not in the same component of a sexual network cannot infect each other with sexually transmitted diseases. A few simple components occur with some frequency in Jefferson High. For example, the simple dyadic structure (two individuals whose only partnership is with each other) occurs 63 times at Jefferson. Thus 126 students are involved in isolated dyadic relations. It is important to note, however, that far more than 126 students at Jefferson report only one relationship; many of the more complex components also include students with only one partner. However, the partners of these students have other partners. This illustrates the importance of collecting data extending beyond ego-centric networks, for it is only by learning directly about the behavior of partners' partners that we can map the structure of connectivity through which disease must flow.[22]

Components involving three students are also fairly prevalent at Jefferson. Triads composed of one male and two females occur 12 times, and triads composed of one female and two males occur nine times. All told, a total of 189 students at Jefferson (35% of the romantically active students) are embedded in sexual and romantic network components containing three or fewer students. There are very few components of intermediate size (4–15 students).

The most striking feature of the network is the existence of a very large component involving 52% ($n = 288$) of the romantically involved students at Jefferson. While this large component involves the vast majority of individuals with multiple partners, it has numerous short branches. Further, it is very broad: the two most distant individuals are 37 steps apart. Most surprising, it is characterized by the almost complete absence of short cycles. Thus the network closely approximates a chainlike spanning tree.

The size of the large component of connected nodes identifies the worst-case scenario for potential disease diffusion within the population. While one-third of all students are embedded in small, disjoint dyads and triads,

[21] A number of readers have asked whether the observed network is a scale-free network with a power law distribution of degree. See app. fig. C1 for an illustration of our response.

[22] Even if each respondent had simply reported on their number of partners in the past 18 months, without their having selected them from a roster of possible partners, we would not have been able to generate this structure, for we would not have known how—or whether—these partnerships connected into a macrostructure. Attribute-based matching schemes are similarly unable to specify the interrelationship of partners' partners.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

in an 18-month period more than 50% of the students at Jefferson were chained together through romantic and sexual relationships that could have involved the exchange of fluids. Recall that there are many individuals at the end of small branches in the large component who have only one partner. While these adolescents have only had one partner, their risk for contracting an STD may be significantly greater than an individual with multiple partners who is embedded in a smaller, disjoint component. Consequently, STD risk is not simply a matter of number of partners.

While it is reasonable to think that an individual might have some sense of their own partners' relationships, the structure of the larger components, and certainly the largest component, is not likely to be visible, or meaningful, to the students at Jefferson. These structures reflect relationships that may be long over, and they link individuals together in chains far too long to be the subject of even the most intense gossip and scrutiny. Nevertheless, they are real: like social facts, they are invisible yet consequential macrostructures that arise as the product of individual agency.

Temporal Unfolding

Figure 2 depicts the direct relationship structure linking individuals together. Disease transmission, however, rests on temporally ordered relationships, and these determine the indirect pathways that can put individuals at risk for disease. Thus if A and B are partners at time 1, and B and C are partners at time 2, from a viral or bacterial perspective a meaningful directed path with the capacity to transmit disease exists between A and C. In contrast, given this pattern of relationships, disease cannot flow from C to A. Taking into account data describing the temporal ordering of relationships, figure 3 reports all indirect and direct ties that could potentially transmit disease within the major component of the Jefferson network.[23] Note that compared to the direct graph in figure 2, the indirect graph is quite dense and contains many regions with interacting adolescents. As a comparison, consider figure 4, the graph of a simulated network containing the *fewest* possible indirect relations derived from the original component. This is the minimal arrangement from a disease perspective; the difference in density between figure 3 and figure 4 suggests the extent to which the actual dynamic unfolding of partnerships at Jefferson increases the potential for widespread disease diffusion (Moody 2002).

[23] A "movie" of the Jefferson network unfolding through time is available at http://www.columbia.edu/iserp/people/bearman/chains.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

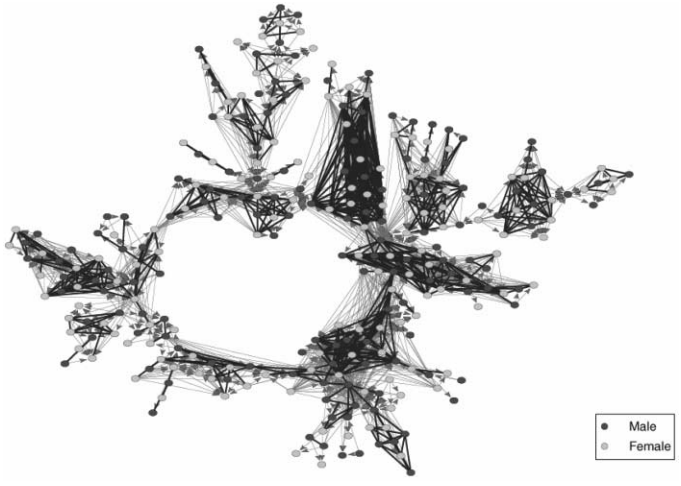

FIG. 3.—Temporally ordered ties in the Jefferson High partnership network

Structural Fragility

Examining the pattern of indirect ties reveals the level of connectivity and redundancy of the network through which disease could travel. While figure 3 reveals the existence of clusters of romantically involved students, it does not reveal how robustly connected these clusters are to one another. In general, structures like spanning trees are considered structurally fragile because the deletion of a single tie or a single node can break a large component into disconnected subgraphs.[24] Consider again the analogy to phone lines: if phone lines are laid out as a spanning tree, a break in the major trunk line separates a single component into two disjoint components and prevents calls from traveling from one component to the other. Engineers protect against such failure by adding lines that build redundancy into the system. The essential structural fragility of spanning trees reveals how subtle changes in local network structure (deleting or adding a relationship, e.g.) can have profound effects at the macrolevel.

Building from the temporally ordered indirect network shown in figure 3, figure 5 reveals how the structure of indirect ties breaks into a set of smaller, mutually reachable sets when cut-points (single pathways between nodes) are eliminated. While each of the remaining smaller components appears to be dense and corelike, simply removing ties at the cut-points fractures the structure into separate components. For sexual

[24] Structural fragility is also referred to as "1-connectedness" in the technical literature on graphs.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

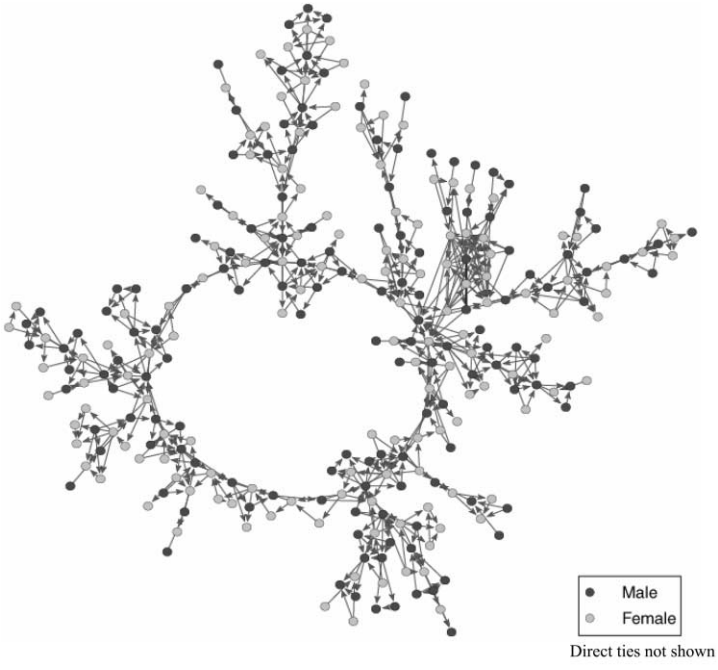

FIG. 4.—Simulated minimal temporal connectivity in the Jefferson High partnership network.

networks, redundant lines provide the foundation for cores, the incubators of epidemics. Thus in a sparse treelike contact structure with many cut-points, failure to transmit disease within a partnership that happens to be a cut-point can break the larger connected components into separate, unconnected subcomponents, thereby fragmenting the potential epidemic.

GENERATING THE STRUCTURE: COMPARISON TO SIMULATED NETWORKS

Data describing the complete structural mapping of a romantic/sexual network in an interacting population has not been previously collected, so there is no obvious baseline against which to evaluate whether what we observe is unusual. Further, the distributional properties of many of the network statistics we are interested in are not well known. Thus while the graph of the observed network at Jefferson appears to stand in clear contrast to the structures implied by most epidemiological models, there is some possibility that it is simply a stochastic realization of one of the random or preferred-mixing models. To test against these alternatives, we

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

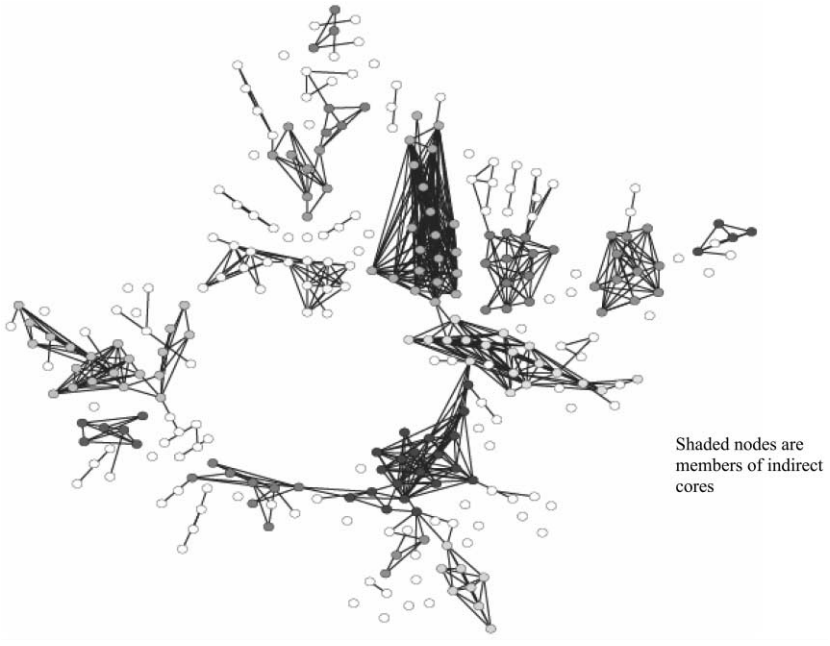

FIG. 5.—Temporally ordered partnerships: cutpoints removed

simulate 1,000 networks consistent with various constraints characteristic of these other models and examine whether the relevant structural characteristics of the Jefferson network are statistically likely, given the distribution of simulated networks.

We begin with the simplest model: Jefferson students select their partners at random.[25] To test this, we simulate 1,000 random networks with the same size and degree as observed in Jefferson,[26] and then we consider where the network at Jefferson falls relative to the distribution of simulated networks.

Figure 6 presents box plots comparing the Jefferson High network to the simulated networks across six measures relevant for STD diffusion dynamics: density at maximum reach, centralization, mean geodesic

[25] The idea seems far-fetched, but, as one reviewer notes, the context is already essentially homophilous with respect to age, education, race, social class, citizenship, religious orientation, and ethnicity—the major determinants of partnership choice for adults. Given this, random choice makes sense as a baseline.

[26] Specifically, we generate 1,000 networks with the same number of nodes and the same distribution of number of partners as observed in the Jefferson net, with ties assigned randomly between nodes. Details about the algorithm used to generate these conditional random graphs can be found in Moody 1998.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

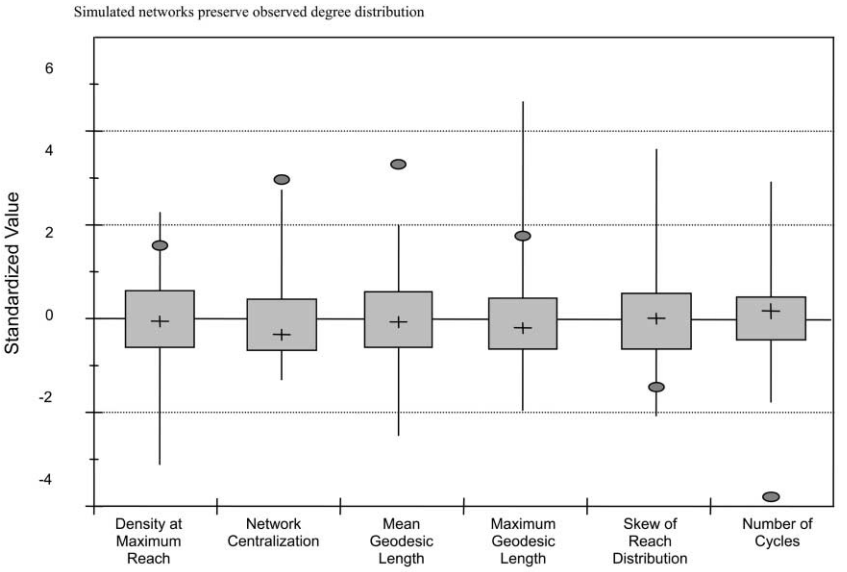

FIG. 6.—Simulated networks, preserving observed degree distribution

length, maximum geodesic length, skew of reach distribution, and number of cycles. The values for each network measure are standardized (mean = 0; SD = 1). The crosshatch within each box plot reports the median value from the simulated nets; the interquartile range is shaded. A dark circle indicates the value we observe for Jefferson High. Not surprisingly, across *all* of these basic measures of network structure, the sexual and romantic network in Jefferson is an outlier relative to the simulated networks generated by random mixing. We discuss each measure in turn.

The first measure, *density at maximum reach*, assesses the extent to which the overall network is connected. Here, we measure the density of the network that arises when ties link all pairs of ever-reachable individuals.[27] All things being equal, heightened connectivity is associated with more efficient disease spread. Compared to the 1,000 simulated networks, the Jefferson network is highly connected. This means that students at Jefferson are more likely to have partners at school who have other partners at school who have other partners at school. The alternative is many dyads, or other small groups of linked adolescents, that are ultimately disjoint from the rest of the population: for example, a core disconnected from a set of smaller components. The highly connected structure at Jefferson, therefore, poses a greater disease risk than would exist if the partnerships were formed at random.

We next consider *network centralization*, a measure of the inequality of the centrality of persons in the network. We calculate centralization using Bonacich's centrality algorithm with a negative beta value (Bonacich 1987). This parameterization gives extra weight to individuals who connect otherwise unconnected individuals and less weight to those whose bridges are redundant to other bridges in the network. Compared to simulated networks with the same degree distribution, the Jefferson network is highly centralized, suggesting that some actors play a central role in linking disjoint clusters.

Continuing to move from left to right, the next box plot compares the *mean geodesic length* of the Jefferson network to the simulated networks. A geodesic is the shortest path between two connected persons in a network: mean geodesic length is the mean of the shortest path between every connected pair in the network. Large geodesics are characteristic of spanning trees, the sparse chainlike structures with few alternate paths

[27] Alternatively, one could represent this measure as the mean number reachable in a network. Likewise, from the same framework one can assess whether the maximum reach of the largest component, in our case involving roughly one-half of all students in the school, is to be expected by chance. Relative to the simulated samples, the largest component is almost 2 SDs greater than expected by a random-mixing model.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

directly connecting persons. In contrast, holding connectivity constant, networks containing redundant links between actors (cycles, a core, or starlike structures) will have smaller mean geodesic lengths. With respect to STD diffusion, the absence of redundancy places pressure on the values of the β (probability of transmission given contact) and D (duration of infectiousness) parameters discussed previously. If β and/or D are low, spanning trees are inefficient structures for diffusion of STDs. Compared to the simulated networks, the network at Jefferson High has very long mean geodesics. This is the result of the extremely large component and the overall absence of "short-cuts," or redundant ties, within the large component.

While the Jefferson network is highly connected, this connectivity is the result of very long chains. It follows that as the mean geodesic length is large, that the *maximum geodesic length*—a measure that captures the number of steps between the most distant pair of connected persons—will also be large. And compared to the simulated networks, the most distant pairs of connected individuals in Jefferson are quite distant from one another. In fact, they are not likely to know that they are involved in the same romantic web, which exists as a social fact beyond the reach of ordinary cognition.

By definition, every person in both the simulated and observed networks is connected to at least one other person (their romantic relationship partner). In addition to their own direct relations, however, they may be indirectly connected to others through the relations of their partners and their partners' partners. Extending this logic, we can calculate for each actor an individual-level measure of the number of "reachable" alters in the network. We then can consider the *skew of the reach distribution* — how unequally the number of reachable partners is distributed across the population. If most of the population were in isolated dyads, the distribution would show a strong positive skew, and the structure would contain few efficient pathways for disease transmission. In contrast, a network that includes a very large component would show a strong negative skew. This is the case in Jefferson. Negatively skewed reach distributions are a trace of contact structures with heightened potential for disease spread.

Among the most structurally characteristic feature of the graph of the Jefferson network is the pronounced absence of *short cycles*. The absence of short cycles guarantees that we do not observe a densely interconnected core that has the capacity to function as a disease reservoir. In comparison to the simulated networks, the romantic and sexual network at Jefferson is characterized by significantly *fewer* cycles than occur when partnerships are chosen randomly. Consequently, STD models that assume a core or inverse core structure are not appropriate here. Using such models in contexts such as Jefferson could result in underestimation of the potential

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

for disease spread—especially if β or D is moderately high, as is the case when treatment rates are low or asymptomatic cases are frequent.

Compared to randomly simulated networks of similar size and degree, the empirical sexual network we observe is quite distinctive. The Jefferson network is dominated by an extraordinarily large component that connects more than half of all the students who are romantically and sexually active in the school. Yet while this component ties individuals together into long chains of potential infectivity, it is extremely fragile. This fragility is largely due to the striking absence of cycles (redundant paths) in the large component.

PREFERENCES FOR PARTNERS

Because the spanning tree structure we observe is extremely unlikely to be the result of random mixing, some other set of processes governing partnership selection must account for it. It is obvious that, when individuals choose partners, they do not base their choice on its contribution to the global macrostructure. Put most starkly, adolescents do not account for their partner choice by saying, "By selecting this partner, I maximize the probability of inducing a spanning tree." First, they cannot see the global structure, and second, they do not care about it. They do care, however, about the more immediate local structure in which the partnership is embedded, and they care about the attributes their potential partner has.

One possibility is that there is a simple micropreference governing choice that, if followed by most individuals, would naturally produce a spanning tree. This is the solution we ultimately consider. We propose a specific rule that, if followed, induces the macrostructure we observe, given the conditions of partnership preference in Jefferson. Later in this article we provide reasons for thinking that this preference is enacted, even if adolescents do not articulate it. Getting to this point requires examining the empirical determinants of partnership choice at Jefferson, which we consider immediately below.

Attribute-Based Selection Preferences

Everyday experience, a cursory glance at personal advertisements in the classified section of any newspaper, a brief inquiry into the underlying logic of dating or matchmaking services, and a wide body of research all suggest that individuals select partners on the basis of characteristics, and that persons tend to prefer partners who are similar to them. While the number of attributes and behaviors that could provide a foundation for

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

preferential partnership selection is enormous, in table 2 we report the level of homophily across a set of attributes and behaviors that might reasonably be expected to govern partnership formation among adolescents.[28] To assess the extent of homophily on selected attributes within romantic partnerships in Jefferson, we generated 500 permutations of the attribute distance/matching matrix with the romantic relation matrix, and then we used QAP to evaluate the difference in attribute means between actual romantic pairs and the randomly simulated partnerships.[29] For continuous variables, the test statistic compares the mean of the difference in the absolute value of the attribute measure for romantic pairs with the mean of the difference between the randomly assigned pairs.[30] Thus, for example, the 0.367 value for grade means that, on average, romantic pairs are about a third of a grade closer to each other than are randomly assigned pairs. For categorical variables (i.e., smoking), the test compares matching scores between real and randomly assigned pairs ($X_{ij} = 1$ if $x_i = x_j$ where X is the matching indicator, x is the attribute, and i and j are the members of the pair). The difference formula is then the proportion of nonromantic pairs that match minus the proportion of romantic pairs that match. Thus the value of $-.11$ for smoking means that the proportion of similar-smoking-status romantic pairs is .11 larger than that for randomly assigned pairs.

Table 2 demonstrates clear evidence of homophily in romantic partnerships.[31] Adolescents at Jefferson tend to select partners with similar socioeconomic status, grade point average, college plans, attachment to school, trouble in school, drinking behavior, IQ, and grade. With respect to categorical attributes, partners tend to be similar in terms of sexual experience, suspension from school, and smoking. Less important is religious denomination. Evidently, students who smoke prefer other students who smoke. Alternatively, students who smoke induce smoking in their partners, perhaps because only smokers can tolerate kissing smokers.

[28] Obviously, judgment is required here. Adolescents may select partners on the basis of unobserved characteristics (or unobserved to us) that vary across individuals in a completely unsystematic way. This is, in one sense, what the idea of romantic love suggests. Our strategy is to identify a set of characteristics that are observable, common, and have face validity as salient attributes. We consider homophily on these, and then simulate the global structure that would arise, should these elements provide the basis for choice. We do not include one of the most salient attributes for partnership choice, race, since Jefferson is all white.

[29] We thank an *AJS* reviewer for suggesting a nonparametric analysis strategy (QAP) in this context.

[30] Specifically, mean[abs$(X_i - X_j)|ij = 0$] − mean[abs$(X_i - X_j)|ij = 1$].

[31] It is important to recall that the context is already quite homogeneous. So, among white students of roughly the same social class, ethnicity, citizenship, and so on, these are the salient determinants of partnership choice.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

TABLE 2
HOMOPHILY IN STUDENT PAIRS

VARIABLE	QAP MEAN DIFFERENCE[a]	
	Full Network	Cross-Sex Only
Family SES	.299***	.295***
Grade	.331***	.367***
GPA	.096**	.102***
Expect to graduate college	.202***	.222***
School attachment	.118***	.132***
Trouble in school	.029	.019
Gets drunk	.180***	.195***
Delinquency[b]	−.058	−.070
Hours watching TV	−.149	−.027
Religiosity (praying)	−.006	−.012
Popularity (in-degree)	−.377*	−.211
Self-esteem	.004	.008
Autonomy	.008	.002
Expect to get HIV	.003	−.007
Expect to marry by 25	.025	.020
Attractiveness	.013	.047
Vocabulary (AH_PVT)	1.508***	1.671***
Religion	−.034*	−.043*
Sexually active	−.100***	−.124***
Smoking	−.087***	−.110***
School suspension	−.028	−.066**
Tattoo	−.003	−.016

[a] Significance reflects exact *P*-test comparison to 500 permutations of the attribute distance/matching matrix with the romantic relation matrix.
[b] Delinquency is standardized by gender and age.
* $P < .05$.
** $P < .01$.
*** $P < .001$.

While homophily is strong, the preference for similarity does not extend to all characteristics, most obviously sex and age. Almost every single reported romantic relationship at Jefferson is a cross-sex relationship, and as is true in most high schools, girls at Jefferson tend to be involved with older boys. Ninth grade girls tend to be in relationships with ninth and tenth grade boys, tenth grade girls with boys in the tenth and eleventh grades, and so on. Among all partnerships involving Jefferson students, we observe a mean grade difference of .9, less than expected if relationships were formed independent of age (mean difference = 1.23 in the randomly assigned pairs), but evidence of a female preference for older boys (or male preference for younger girls).[32]

[32] Recall that not all partners, sexual or romantic, are drawn from school. On average, out-of-school partners were 3.21 years older than the respondent at the start of their

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

Homophily in Partnership Experience

Even given these revealed preferences for attributes, adolescents have a great deal of leeway in terms of selecting potential romantic partners. Among adults, we know that experienced partners prefer experienced partners (homophily on experience), a preference that can give rise to cores (Laumann et al. 1994). Visual inspection of the graph shown in figure 2 suggests that many of the differences between the Jefferson sexual and romantic network and the simulated networks may be the result of the large number of isolated dyads we observed in Jefferson. Thus we ask whether the single large component involving half of all students is a mathematical byproduct of homophily on this one partnership characteristic: the number of previous partners an individual has had. If a majority of the individuals with only one partner are involved with individuals who also have only one partner, it would follow that those with multiple partners are constrained to be involved with persons who have also had previous partners. The catenation of these individuals should, all things being equal, generate large interconnected components.

To test this idea, we again simulate 1,000 networks with fixed size and degree distribution, this time removing the 63 isolated dyads (involving the 126 persons whose single partner has only a single partner) and prohibiting the creation of new isolated dyads. We then compare the large component from the Jefferson network to the structural characteristics of network simulated with the prohibition against isolated dyads. Adding this single additional constraint has a stunning impact on the structure of our simulated networks. Specifically, the mean size of the largest component in the simulated networks is now very close to the size of the large component in the Jefferson network (mean of 283 nodes vs. 288 nodes in Jefferson). As suspected, homophily in partnership selection among less experienced partners (those with only a single romantic involvement) *provides an efficient micromechanism for the generation of a large component*. Thus homophily on experience is a key element in generating the structure we observe.[33]

---

relationship, although we observe a pronounced skew in the age-difference distribution. For example, one out-of-school sexual partner of two girls was 39 years old. Aside from a few exceptions like this, most of the students involved in out-of-school relationships have partners slightly older than themselves. Together with other evidence in the survey describing where the respondents met their partners, we conclude that many of these out-of-school partners attended Jefferson prior to our survey, and most continue to live in the same neighborhood.

[33] We have also estimated p^* models that simultaneously evaluate the effect of homophily and structural characteristics on the likelihood of a graph. Because we are skeptical about the interpretability of parameters from p^* models of our data, we do not discuss them in the text. However, we report and discuss these models in app. B.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

To consider network features other than component size, we again compare the simulated networks with the Jefferson network across the set of network measures salient for disease diffusion previously discussed in figure 6. These results are shown in figure 7.

As before, values are standardized (mean $= 0$; SD $= 1$). The crosshatch within each box plot reports the median value, and the inter-quartile range is shaded. Dark circles indicate the values we observe for Jefferson. Across all six network measures, the structure of the Jefferson sexual network remains significantly different than expected, although less so than under the less constrained simulations, where we did not prohibit isolated dyads. The improved fit results from smaller variances in the simulated networks.

Although homophily on experience appears to account for the size of the largest component, the structural characteristics of the observed Jefferson network are still unusual relative to the simulated networks. Thus this micromechanism is not sufficient to reproduce the structural properties of the observed network. The main differences between the simulated and real networks, in mean geodesic length, network centralization, reach, and skew of maximum reach, are the product of the absence of cycles. Thus while preferential selection on partnership experience level provides an efficient foundation for generating large components in adolescent sexual networks, it fails to generate a spanning tree. Among romantically active students, random mixing produces more redundant ties than exist in Jefferson. As a consequence, the simulated networks reveal core structures rather than spanning trees.

UNCOVERING GOVERNING NORMS

Our analyses thus far demonstrate that the macrolevel network structure at Jefferson is neither the simple product of random mixing nor of individual preferences for partners with particular attributes. Because we find many cycles of length 4 in the simulated networks, but few in Jefferson, we believe that there must be a prohibition against partnerships that involve the creation of short cycles.

We adopt a new strategy to investigate just how unusual short cycles are at Jefferson. Earlier we showed that while spanning trees may be efficient for disease transmission, they are structurally fragile. Whereas our investigation of structural fragility was based on the consequences of *removing* relationships from the graph, we now consider the effects of *random rewiring* of the network. That is, we randomly reassign partnerships from one pair of nodes to another pair. Since the new partnerships we introduce are formed at random, they are insensitive to any existing norms or preferences that may govern partner choice at Jefferson. By

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

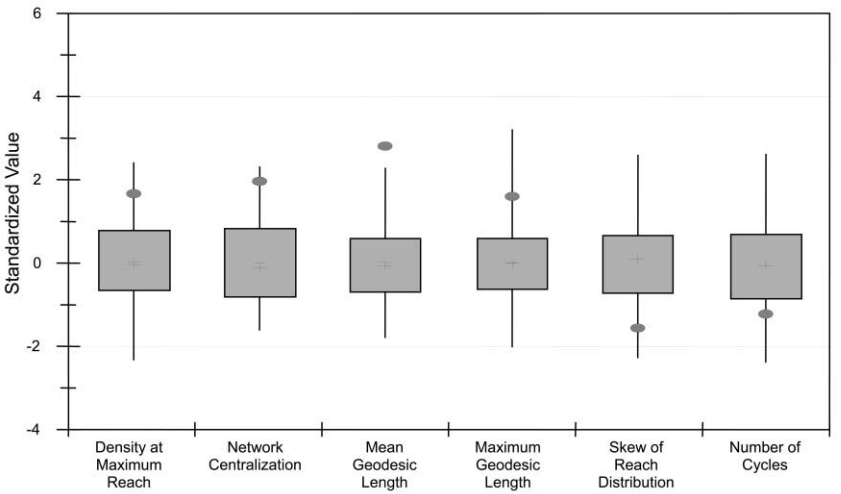

FIG. 7.—Simulated networks, preserving observed degree and isolated dyad distribution

analogy, consider the existence of an incest taboo that restricts available partners to those who are a culturally agreed upon distance from ego. "Rewiring" the resulting marriage graph means that some relations prohibited by this rule will be added to the network, and therefore some structural properties of the new graph may no longer match the original data structure. Structural features of the rewired graph that deviate from the original graph help us identify behavioral rules that govern that specific parameter.

To rewire the empirically observed graph, we select 5% of the relationships at random and reassign them conditional only on the degree distribution of the original graph. In this way, we ensure that individuals with many partners continue to have many partners and that individuals with few partners do not suddenly gain partners. Table 3 reports the effect of rewiring and is based on a comparison of the observed network with 1,000 rewired graphs.

Compared with our earlier simulations, the rewired graphs are quite similar to the observed network. This is what we would expect, since we change only 5% of the ties at random, while holding the distribution constant. Consequently, all of the network centralization measures are fit well, as are the reach measures. The difference between the observed number of components in Jefferson and those arising from the simulations is trivial. *The only statistic that is fit poorly is the number of cycles.* The rewired networks have almost twice as many cycles as are observed in Jefferson. Since we observe a spanning tree in the Jefferson network, it is not surprising that rewiring produces redundant ties, which appear here as cycles. Thus rewiring isolates the single structural feature we have to account for—in this case, the absence of cycles. Thus the only puzzle is, Why are they absent?

The Basis for a Spanning Tree Structure: Unarticulated Partnership Prohibitions

To explain why cycles are absent at Jefferson, recall that spanning trees are theoretically produced by negative proscriptions. What kinds of relationships are prohibited? The simple answer is that the prohibited relationships are those that induce short cycles. The *smallest possible heterosexual cycle* has a length of 4. Consider four individuals, Bob, Carol, Ted, and Alice. Imagine that Bob and Carol were once partners, but that Carol left Bob for Ted. Further imagine that Ted and Alice were partners, but that Ted dumped Alice for Carol. Should Bob and Alice date? From Bob's perspective, Alice was his former partner's current partner's partner, or the former "lover" of his former girlfriend's current lover. Alice looks at Bob with the same lens. Her former boyfriend is dating the girl

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

TABLE 3
ROBUSTNESS OF THE OBSERVED STATISTICS TO RANDOM REWIRING

VARIABLE	OBSERVED NETWORK*		RANDOM REWIRE DISTRIBUTION	
	Coefficient	*P*	Coefficient	SD
Density at maximum reach ...	.42	.34	.43	.06
Network centralization	.024	.23	.023	.003
Mean geodesic length	15.94	.10	13.8	1.71
Maximum geodesic length	37	.30	34.9	5.52
Reach distribution skew	−.61	.34	−.67	.21
N of cycles	5	.04	9.03	2.26

NOTE.—*N* of rewired nets = 1,000.
* Excludes isolated dyads.

who left Bob.[34] These scenarios can be summarized by a graph, as in figure 8, where lines indicate a relationship between nodes (here, persons), yielding a potential cycle of length 4.

Using the simulation strategy introduced earlier, we can operationalize a normative rule that persons do not date the former (or current) partner of their former (or current) partner by prohibiting all cycles of length 4. We simulate 1,000 random networks, this time conditional on the following constraints: fixed degree distributions matching those observed at Jefferson, no isolated dyads, and a single parameter that prohibits cycles of length 4. The question is whether this set of constraints generates graphs with structural features similar to those observed at Jefferson. Figure 9 shows that they do: on all the structural parameters we consider, the Jefferson network is quite close to the central tendency of the distributions generated by the simulated networks.

Comparison of the internal structure of these random networks and the observed Jefferson network shows that they are essentially isomorphic. This similarity is illustrated in figure 10, which shows graphs of the largest components from four randomly selected networks simulated by this model. One immediately sees network structures strikingly similar to the structure observed in Jefferson. Given fixed degree and homophily in experience, the sufficient condition for generating a spanning tree is the prohibition against cycles of length 4. Such a prohibition may operate in Jefferson.

[34] Here preferences for partner attributes break down: if Carol is attracted to both Bob and Ted, then they must be similar with respect to attributes, and yet clearly Carol and Alice are not equivalent substitutes from Bob's perspective.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

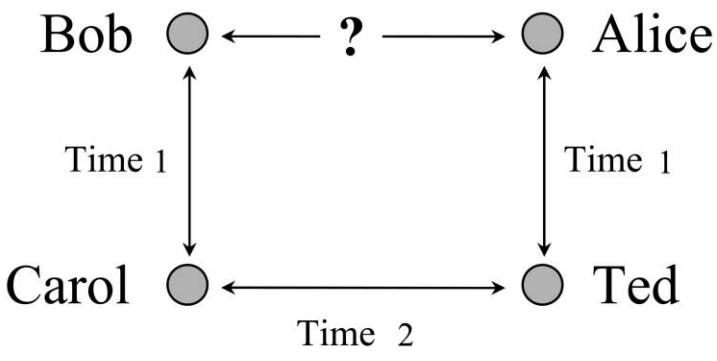

FIG. 8.—Hypothetical cycle of length 4

Status Dislocation and Closeness

Given the conditions of homophily described previously, figures 9 and 10 show that a simple rule—the prohibition against dating (from a female perspective) one's old boyfriend's current girlfriend's old boyfriend—accounts for the structure of the romantic network at Jefferson. Why might this negative proscription operate in a medium-sized community of essentially homogenous adolescents?

The explanation we offer only makes sense for short cycles. From the perspective of males or females (and independent of the pattern of "rejection"), a relationship that completes a cycle of length 4 can be thought of as a "seconds partnership," and therefore involves a public loss of status.[35] Most adolescents would probably stare blankly at the researcher who asked boys: Is there a prohibition in your school against being in a relationship with your old girlfriend's current boyfriend's old girlfriend? It is a mouthful, but it makes intuitive sense. Like adults, adolescents choose partners with purpose from the pool of eligible partners. But beyond preferences for some types of partners over others—for example, preferences for partners interested in athletics, who do not smoke, or who will skip school to have more fun—adolescents prefer partners who will not cause them to lose status in the eyes of their peers. In the same way that high-status students avoid relationships with low-status students, by selecting partners on the basis of the characteristics that have resonance

[35] The status-loss hypothesis competes with other potential micromechanisms, e.g., "jealousy" or the avoidance of too much "closeness," a sentiment perhaps best described unscientifically as the "yuck factor." The status-loss hypothesis involves significant scope limitations: namely, status loss is limited to contexts where actors by virtue of their relational density can watch each other relatively closely. By contrast, the "yuck factor"—which is essentially individualized—could operate in more diffuse contexts.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

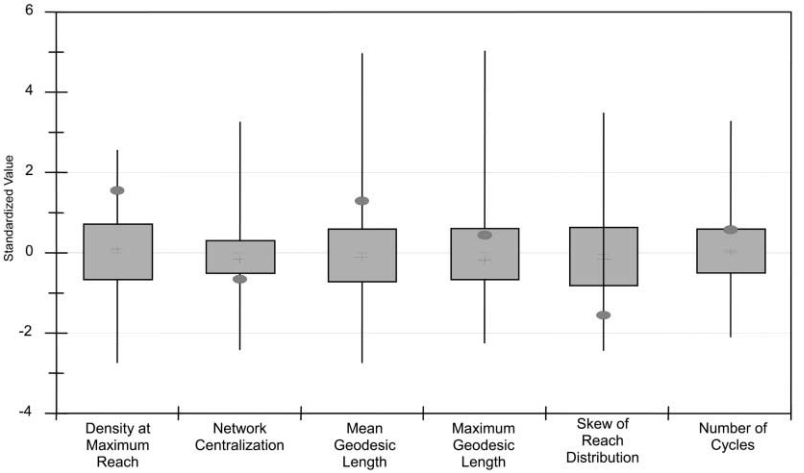

FIG. 9.—Simulated networks preserve observed degree, isolated dyad distribution, and four-cycle constraint

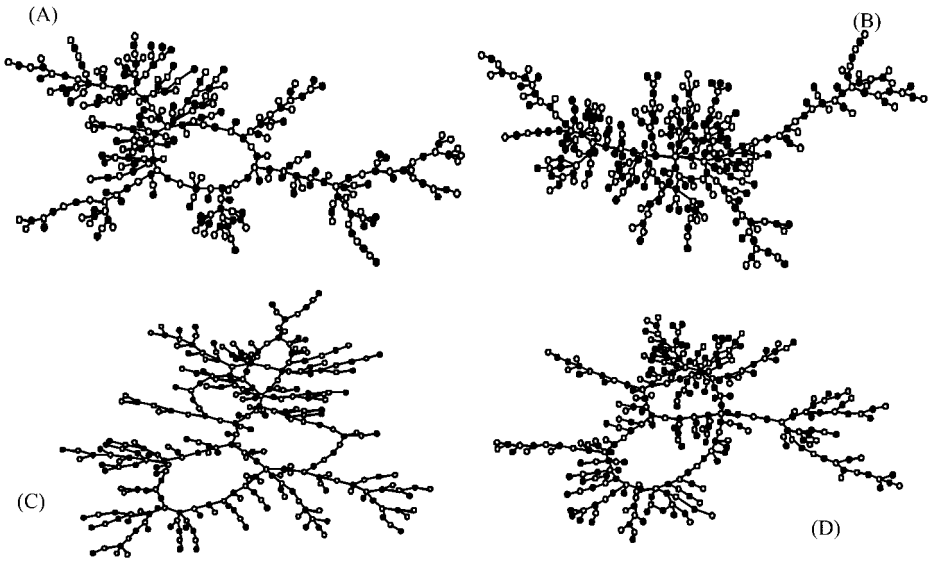

FIG. 10.—Simulated networks preserve observed degree, isolated dyad distribution, and cycle constraint

for the local determination of prestige, students avoid relationships whose *structure* would lower their status in the eyes of their peers. In a large and essentially homogenous school like Jefferson, the pool of potential partners with the "right" mix of attributes is relatively large, so students can fairly easily avoid taking "seconds" and still preserve their basic attribute or experience preferences. More generally, in intact communities where observation of temporally proximate partnerships is possible, we should expect to see successful avoidance of relationships that complete cycles of length 4. Such avoidance should not, however, extend to larger cycles, since larger cycles typically involve relationships in the distant romantic past that cannot be systematically observed.[36]

For adolescents, the consequence of this prohibition is of little interest: what concerns them is avoiding status loss. But from the perspective of those interested in understanding the determinants of disease diffusion, the significance of a norm against relationships that complete short cycles is profound. The structural impact of the norm is that it induces a spanning tree, as versus a structure characterized by many densely connected pockets of activity (i.e., a core structure). As a consequence this prohibition, combined with existing homophily preferences, both shapes the potential for local disease diffusion and affects what social policy interventions will be effective at stemming disease spread.

DISCUSSION

Disease diffusion is widespread among adolescent populations. The standard models that epidemiologists use to describe the dynamics of diffusion carry implicit ideas about the contact structure through which disease travels. These ideas are associated with distinct structural features of sexual networks. The most critical feature in STD epidemiology is the idea of a core, which is associated with cycles in networks. Moody et al. (2003) have demonstrated that very low average degree networks can give rise to densely interconnected cores, characterized by high cyclicity. In our data, we find that this key structural feature is largely absent. We have proposed a reason for its absence, specifically a norm against seconds partnerships. From this norm, combined with basic homophily preferences, we generate networks that are structurally isomorphic to the one we observe empirically. This suggests that in adolescent society—where

[36] In a sense the prohibition against cycles of length 4 suggests shifting from the cross-sectional perspective considered in fig. 2 and focusing instead on the temporally ordered graph shown in fig. 3, since a cycle unfolds over time. The picture is made somewhat more complex by instances in which relationships are concurrent. While not uncommon at Jefferson, concurrency appears to be more common among adults than adolescents.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

partner choice is salient for local status—it seems reasonable to think that such a rule operates.[37]

Nonetheless, the scope conditions for this article are implied in the central finding and the mechanism we claim accounts for it. Specifically, our mechanism presumes that actors can watch each other, that they are capable of recording immediately prior partnerships, and that they are susceptible to collective assessment of their personal choices. One can only fear losing status in the eyes of others if the others are watching and if one cares about their assessments.

These conditions may be absent for adults who are embedded in worlds larger and more disjoint than adolescents. More than adolescents, adults may be capable of segregating audiences across the various settings in which they are embedded (work, leisure, play, school, etc.) and are therefore less subject to the scrutiny and sanctioning of their peers than are adolescents. While entering into a partnership that completes a cycle of length 4 may result in a loss of face for an adult, it is more likely that among adults such cycles are generated without anyone ever knowing. This is unlikely in a high school, where much social energy is devoted to understanding who is going out with whom.[38] All this suggests we would be less likely to observe spanning trees among adult populations than among adolescents.

In theory, spanning trees are among the most efficient structures for diffusion since the absence of redundant lines maximizes reach at lowest density. Yet their efficiency is counteracted by their fragility: spanning trees are highly susceptible to breaks in transmission. Electric provision systems would be set up as spanning trees if service providers did not worry about failing to deliver power to some customers. But since they worry about small breaks in the line, they establish more densely connected power grids. It follows that for highly infectious diseases with long

[37] We note again that not all of the sexual partnerships in Jefferson are directed toward other students. It is possible that the spanning tree structure we observe is a by-product of missing data on the prior partnerships of the out-of-school partners. We consider this unlikely for three reasons. First, adolescents with any out-of-school partners are disproportionately older and more likely to be female than are those with only in-school partnerships. Second, those with out-of-school partners are less likely than those with in-school partners to have *other* in-school partnerships. And third, analysis of temporal sequencing indicate that most out-of-school partnerships are temporally subsequent to in-school partnerships, should there be any. Consequently, their impact on the structure of the observed network is necessarily modest, since if present, they do not link (from a viral/bacterial perspective) nodes in the in-school graph we consider.

[38] Because out-of-school partnerships tend to be more privatized—i.e., involve seeing less of friends—the scope conditions are relevant here as well, as privatized relationships are less likely to be observed and hence could, theoretically, not be subject to the four-cycle constraint.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

periods of infectivity, transmission is also quite efficient under a spanning tree. Yet if the duration of infectivity is short, or if the disease is not particularly infectious, the probability of transmission within any given partnership is low. From the perspective of disease spread, failure to transmit disease within a given partnership is effectively a structural break in the network (Watts 2003). Since the natural infectivity and duration of infectiousness varies across STDs, we believe the most effective strategy for reducing disease diffusion rests on creating structural breaks.

We might then ask a new question: What kinds of policy intervention will be most effective at *inducing* structural breaks in the sexual networks of adolescents? Here the answer is exceedingly simple. Assume that some proportion of actors who are "reached" through an intervention decide to change their behavior. Under core and inverse core structures, it matters enormously *which* actors are reached, while under a spanning tree structure the key is not so much *which* actors are reached, just that some are. This is because given the dynamic tendency for unconnected dyads and triads to attach to the main component, the structure is equally sensitive to a break (failure to transmit disease) at any site in the graph. In this way, relatively low levels of behavior change—even by low-risk actors, who are perhaps the easiest to influence—can easily break a spanning tree network into small disconnected components, thereby fragmenting the epidemic and radically limiting its scope. Obviously, similarly low levels of change in cores will have little impact.

This example highlights how having an accurate sense of the real structure of a network matters for the effectiveness of an intervention. If cores exist, one develops interventions that target core members. But if there are no actual structural cores, interventions targeted primarily at high-risk individuals will do less to stem the overall spread of disease than will broadcast interventions directed toward all actors. Ironically, early HIV prevention strategies that utilized broadcast diffusion techniques may have been more effective at reducing overall incidence of disease than more recent interventions focused on isolating those seen as being at a higher risk for infection (though these targeted interventions may reduce risk of disease acquisition at the individual level).

Epidemiologists, unable to observe or measure directly the structure of sexual networks, have tended to latch onto a single idea: specifically, the idea that the number of partners matters for STD diffusion dynamics. If, as their models assume, the real contact structures are well approximated by core or inverse core network structures, degree distributions are meaningful, and the number of partners will be a key parameter. But this need not be the case theoretically, and, as we empirically show, it is not likely the case for adolescents. The fact that the relevant contact structure is a spanning tree explains why the rates of bacterial STDs have been so high

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

among adolescents in the past decade, and why most social policy, which focuses on high-risk individuals within the adolescent community, has failed to stem the flood of new infections. Our data suggest that a shift in social policy toward comprehensive STD education for all adolescents, not just those at highest risk, would be significantly more effective than current intervention models.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

APPENDIX A

Description of Individual-Level Measures

TABLE A1
VARIABLE DEFINITIONS

Variable	Description	Range
Family SES	Occupational prestige and educational background	0–10
GPA	Average of grades in English, math, history, and science	4 = A; 0 = F
Expect college graduation	Respondent's sense of how likely he or she is to go to college	5 = likely; 1 = unlikely
School attachment	Mean of (1) Do you feel close to people at your school? (2) Are you happy to be at your school? (3) Do you feel a part of your school?	5 = strongly agree; 1 = strongly disagree
Trouble in school	Mean of four items about relations with teachers, students, and schoolwork	5 = everyday; 0 = never
Drunk	Frequency of getting drunk	6 = almost every day; 0 = never
Delinquency	Delinquency scale: mean of 15 items about theft, property damage, and violence	5 = three or more times; 0 = never
Hours watching TV	"How many hours a week do you watch TV?"	1–99 hours
Religiosity (praying)	"How often do you pray?"	5 = once a day; 1 = never
In-degree*	Number of times respondent is nominated as a friend	0–37
Self-esteem	Mean of 10 items	4 = high; 0 = low

Sexually active	Self-report of having had sexual intercourse	1 = yes; 0 = no
Autonomy	Sum of seven items about respondent's ability to make own decisions	7 = high autonomy; 0 = low autonomy
Expect to get AIDS	"What do you think are the chances you will get HIV or AIDS?"	5 = it will happen; 1 = no chance
Marry by 25	"What do you think are the chances you will be married by age 25?"	5 = it will happen; 1 = no chance
Attractiveness	Interviewer assessment of student's attractiveness	8 = very attractive; 1 = unattractive
AH_PVT	*Add Health* Picture Vocabulary Test	13 – 146
Two biological parents	Respondent lives with two biological parents	1 = yes; 0 = no
Smokes regularly	Respondent smoked at least one cigarette a day for last 30 days	1 = yes; 0 = no
School suspension	"Have you ever received an out-of-school suspension?"	1 = yes; 0 = no
Tattoo	"Do you have a permanent tattoo?"	1 = yes; 0 = no
Grade	Year in school	Seventh to twelfth grade
Major religion	Respondent's religion	None, Protestant, Catholic, other

* Measurement of *in-degree* is derived from the in-school instrument

APPENDIX B

Building Macrostructure from Homophily on Partner Preferences: A *P** Approach

Here we consider p^* estimates for the Jefferson sexual and romantic networks. The p^* modeling framework has been proposed as a statistical modeling tool for network data. Statistical modeling of social networks is difficult, since the basic assumption of case independence is violated. The p^* modeling framework assumes that dependencies in the network can be modeled as a Markov process (Anderson et al. 1999; Wasserman and Pattison 1996), where dependencies are specified based on particular (usually local) network patterns, such as the number of two-stars or complete triads found in the network. When the p^* assumptions are met, pseudolikelihood logit models can be used to estimate parameters and statistical significance by constructing a set of structural change statistics around the dependencies in the model and regressing the presence of a tie on changes in structural parameters due to the presence of that tie. Among the most useful aspects of p^* modeling is the ability to model homophily parameters while simultaneously accounting for (at least some of) the structural dependencies in the network.

While promising, the p^* modeling framework is not uncontroversial, and the practical application of such models is limited. First, small violations of the Markov assumptions can create biased pseudolikelihood estimates, making it difficult to interpret both parameter estimates and significance tests. Second, the structural change variables used in the logit model estimation are often highly collinear, creating parameter instability and wildly fluctuating model fit. This is a problem we encounter with these data, where some of the structural variables correlate at greater than 0.9. Nevertheless, we consider the analysis strategy here.[39] In the tables below, we estimate p^* models for the Jefferson High School romantic network based on homophily parameters and network structure.

[39] One reviewer suggests that we use Snijders' SIENNA algorithm for dynamic p^* modeling (Snijders 1991). The principal rationale for the decision to not use SIENNA is that the program was not designed for a bipartite graph (which our strong heterosexual network gives us). In addition, SIENNA is not designed for graphs of this size. Even limiting the network to just the largest component (which would be selection on the dependent variable) results in a graph of 288 nodes. Finally, the time features of our graph would result in very sparse networks for any given time window, which means the structural parameters we can include would be too thin and estimation unreliable. The key temporal element that one ought to focus on, we believe is the time-dependent nature of cycles. We have examined this issue descriptively and note that there are never any cycles existing at a particular moment in time. As noted in the text, readers interested in viewing the time ordered evolution of the largest component can request a "movie" of the evolution of the network.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

We treat the graph as strictly heterosexual, removing the small number of homosexual relations. We note that while there are few of these, they are important for the observed structure of the graph, since one of these relations is part of the large cycle evident in the center of figure 2. The effects of removing these cases are a reduction from five to three cycles, a slight decrease in the size of the largest component (from 288 to 287 nodes), and an increase in the diameter of the graph (from 37 to 43). We follow standard practice and build models from a simple random graph to more complex structures. We include the four pairwise attribute measures that provide the best model fits (grade in school, smoking status, popularity, and attractiveness) as well as structural change statistics associated with two paths, two stars, three stars, three paths, and four cycles.[40] To fit the model we construct a male-female dyad data set and estimate the probability of a tie conditional on both individual characteristics and the structural change statistics. For the models that include attributes, cases with missing data on any of the attribute measures are dropped from the network. This, of course, changes the network composition.

The overall model fits for each model we estimate are reported in table B1; the parameter estimates of the attribute mixing for model 12, the model that best approximates the observed data, are reported in table B2.

[40] We do not report results from introduction of the four-cycle term. The coefficient, rather than being negative, is positive, which implies a tendency in the network for four cycles to form. This is empirically incorrect. In fact, this result is picking up the three observed four cycles that come from one high-degree male star. As such, the basic "homogeneity" assumption of the model is not met. If one runs a model either (*a*) without this node in the underlying graph or (*b*) with an interaction that separates out cycles through this one actor, the parameter is negative.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

TABLE B1
P* Models of Observed Romantic and Sexual Networks

Model	Description	No. of Parameters	Correlation between Observed and Predicted	−2 Log-Likelihood Score
1	Simple random graph	1	NA	5,765.21
2	fixed row-column effects	446	.056	5,533.38
3	2 + 2-stars	3	.232	5,374.91
4	3 + 3-stars	5	.409	4,617.08
5	4 + 3-path	6	.410	4,617.03
6	5 + isolated dyads	7	.525	4,511.97
7	1, limited to all nonmissing data	1	NA	4,275.61
8	7 + 2-stars (= 3)	3	.224	4,012.52
9	8 + 3-stars (= 4)	5	.401	3,487.85
10	9 + 3-paths (= 5)	6	.401	3,487.53
11	10 + isolated dyads (= 7)	7	.482	3,443.27
12	11 + attributes	11	.500	3,390.79

TABLE B2
PARAMETER FITS

Variable	Estimate
Density	1.020***
2-stars (M→F)	−2.477***
2-stars (F→M)	−3.864***
3-stars (M→F)	.910***
3-stars (F→M)	1.679***
3-paths	.021
Isolated dyads	−.948***
Grade difference	.092*
Attractiveness difference	−.156**
Both smoke	.841***
Popularity difference	−.076*

* $P < .05$.
** $P < .001$.
*** $P < .0001$.

The homophily effects can be seen with attractiveness and smoking. Similarly attractive students (or, similarly unattractive students) are disproportionately drawn into pairs; on the other hand, smokers and nonsmokers do not pair up often. However, the structural coefficients are incoherent, which suggests that while individuals have preferences, almost any structure is available for their expression. The relative independence of preference to structure provides insight into why such preferences can exist in the first place. If preferences could find expression in only a single structure, most people would be without partners. The fluidity of structures relative to preferences allows all partnerships to be the product of choice, even if the choices and preferences that underlie them are, in a fundamental sense, completely arbitrary, or, alternatively, highly individuated.

APPENDIX C

Q&A

Question 1. One reviewer suggests that if sex "has an element of contagion in which only sexually experienced actors recruit new participants," the resulting graph would be a spanning tree. This model would necessarily imply disassortative mixing by degree. Although interesting, such a model does not correspond to the empirical reality we observe. The simplest way to examine this is to consider the degree distribution for our data (table C1) and compare the observed mixing matrix by degree to the expected mixing matrix (table C2).

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

TABLE C1
OBSERVED DEGREE DISTRIBUTION

Degree	Frequency
1	338
2	139
3	59
4	28
5	6
6	1
9	1

TABLE C2
MIXING MATRIX BY DEGREE

Mixing Matrix	1	2	3	4	5+
Observed:					
1	126	101	54	38	19
2	101	78	49	34	16
3	54	49	38	29	7
4	38	34	29	8	3
5+	19	16	7	3	0
Expected:					
1	120	98.9	63.0	39.8	16.0
2	98.9	81.4	51.8	32.8	13.2
3	63.0	51.8	33.0	20.9	8.38
4	39.8	32.8	20.9	13.2	5.31
5+	16.0	13.2	8.38	5.31	2.13

The observed values (table C1) are quite close to the expected values (the lower half of table C2). Consequently, there is little evidence for *disassortative* mixing. A simple measure of *assortative* mixing would be the correlation between actor i's degree and actor j's degree. In our network, the correlation is $-.03$ ($P = 0.53$), when using the full degree distribution, and $-.02$ ($P = 0.68$) for the collapsed data reported above. Here also, there is no evidence for disassortative mixing or assortative mixing; in short, while interesting, there is no evidence that a contagion model fits the data we observe.

Question 2. A number of readers have asked whether the observed network is a scale-free network with a power law distribution of degree. We have considered this in a log-log plot in figure C1. The slope of the fitted line (the exponent for the power law) has a value of -2.9. While a casual glance suggests that the fitted line mirrors the degree distribution,

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

following Jones and Handcock (2003), we consider it unlikely that the observed degree distribution is in fact scale free.

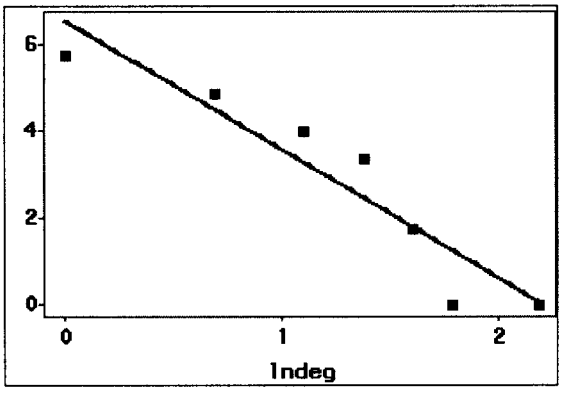

FIG. C1.—The observed degree distribution is not scale free

REFERENCES

Anderson, C. J., S. Wasserman, and B. Crouch. 1999. "A P* Primer: Logit Models for Social Networks." *Social Networks* 21:37–66.

Anderson, R. M., and R. M. May. 1991. *Infectious Diseases of Humans' Dynamics and Control*. New York: Oxford University Press.

Aral, S. O. 2000. "Behavioral Aspects of Sexually Transmitted Diseases: Core Groups and Bridge Populations." *Sexually Transmitted Diseases* 27:327–28.

Aral S. O., J. P. Hughes, B. Stoner, W. Whittington, H. H. Handsfield , R. M. Anderson, and K. K. Holmes. 1999. "Sexual Mixing Patterns in the Spread of Gonococcal and Chlamydial Infections." *American Journal of Public Health* 89 (6): 825–33.

Barabasi, Albert-Laszlo, and Reka Albert. 1999 "Emergence of Scaling in Random Networks." *Science* 286:509–12.

Bearman Peter S., and Hannah Brückner. 1999. *Peer Effects on Adolescent Girls' Sexual Debut and Pregnancy*. Washington D.C.: National Campaign to Prevent Teen Pregnancy.

Bearman Peter S., Henning Hillmann, and Hannah Brückner. 2001. "PAIRS: Adolescent Relationship Dynamics." Manuscript. Columbia University, Institute for Social and Economic Research and Policy.

Bonacich, Phillip. 1987. "Power and Centrality: A Family of Measures." *American Journal of Sociology* 92:1170–82.

CDC (Centers for Disease Control and Prevention). 2000. *Tracking the Hidden Epidemics: Trends in STDs in the United States*. Atlanta, Georgia.

Conover, W. J. 1980. *Practical Nonparametric Statistics*, 2d ed. New York: John Wiley.

Dezso, Zoltan, and Barabasi, Albert-Laszlo. 2002. "Halting Viruses in Scale-Free Networks." *Physical Review E* 65:055103.

Feld, S. L. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86:1015–35.

Friedman, S. R., A. Neaigus, B. Jose, R. Curtis, M. Goldstein, G. Ildefonso, R. B. Rothenberg, and D. C. Des Jarlais. 1997. "Sociometric Risk Networks and HIV Risk." *American Journal of Public Health* 87:1289–96.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

Garnett G. P., J. P. Hughes, R. M. Anderson, B. P. Stoner, S. O. Aral, W. L. Whittington, H. H. Handsfield, and K. K. Holmes. 1996. "Sexual Mixing Patterns of Patients Attending Sexually Transmitted Disease Clinics." *Sexually Transmitted Diseases* 23 (3): 248–57.

Gorbach P. M., H. Sopheab, T. Phalla, H. B. Leng, S. Mills, A. Bennett, K. K. Holmes. 2000. "Sexual Bridging by Cambodian Men: Potential Importance for General Population Spread of STD and HIV Epidemics." *Sexually Transmitted Diseases* 27: 320–26.

Hage, P., and F. Harary. 1983. *Structural Models in Anthropology*. Cambridge: Cambridge University Press.

———. 1996. *Island Networks: Communication, Kinship, and Classification Structure in Oceania*. New York: Cambridge University Press.

Hethcote, H. 2000. "The Mathematics of Infectious Diseases." *SIAM Review* 42: 599–653.

Hethcote, H. W., and J. W. Van Ark. 1987. "Epidemiological Models for Heterogeneous Populations: Proportionate Mixing, Parameter Estimation, and Immunization Programs." *Mathematical Biosciences* 84:85–118.

Hethcote, H. W., and P. van den Driessche. 2000. "Two SIS Epidemiologic Models with Delays." *Journal of Mathematical Biology* 40:3–26.

Hethcote, H., and J. A. Yorke. 1984*a*. "Gonorrhea Modeling: A Comparison of Central Methods." *Mathematical Biosciences* 58:93–109.

———. 1984*b*. *Gonorrhea Transmission Dynamics and Control*. Berlin: Springer Verlag.

Holmes, K. K., P. A. Mardh, P. F. Sparling, S. M. Lemon, W. E. Stamm, P. Piot, J. N. Wasserheit, eds. 1999. *Sexually Transmitted Diseases*, 3d ed. New York: McGraw Hill.

Jacquez, J. A., C. P. Simon, J. Koopman, L. Sattenspiel, and T. Perry. 1988. "Modeling and Analyzing HIV Transmission: The Effect of Contact Patterns." *Mathematical Biosciences* 92:119–99.

Jones, J., and M. Handcock. 2003. "Sexual Contacts and Epidemic Thresholds." *Nature* 423: 605–6.

Klovdahl, A. S. 1985. "Social Networks and the Spread of Infectious Diseases: The AIDS Example." *Social Science Medicine* 21:1203–16.

Koopman, J., C. P. Simon, J. A. Jacquez, and T. S. Park. 1989. "Selective Contact within Structured Mixing with an Application to HIV Transmission Risk from Oral and Anal Sex." Pp. 316–49 in *Mathematical and Statistical Approaches to AIDS Epidemiology*, edited by C. Castillo-Chavez. Berlin: Springer-Verlag.

Laumann, Edward O., John H. Gagnon, Robert T. Michael, and Stuart Michaels. 1994. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.

Laumann, Edward O., and Yoosik Youm. 1999. "Racial/Ethnic Differences in the Prevalence of Sexually Transmitted Diseases in the United States: A Network Explanation." *Sexually Transmitted Diseases* 26:250–61.

Moody, James. 1998. "A General Method for Creating Approximate Conditionally Uniform Random Graphs." Paper presented at the Sunbelt Social Networks Conference. Sitges, Spain.

———. 2002. "The Importance of Relationship Timing for Diffusion: Indirect Connectivity and STD Infection Risk." *Social Forces* 81:25–56.

Moody, James, Martina Morris, Jimi Adams, and Mark Handcock. 2003. "Epidemic Potential in Low Degree Networks." Manuscript. Ohio State University.

Morris, Martina. 1993. "Epidemiology and Social Networks: Modeling Structured Diffusion." *Sociological Methods and Research* 22:99–126.

———. 1997. "Sexual Networks and HIV." *AIDS 97: Year in Review* 11:S209–S216.

Morris, Martina, and M. Kretzschmar. 1995. "Concurrent Partnerships and

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions

Transmission Dynamics in Networks." Special issue, Social Networks and Infectious Disease: HIV/AIDS. *Social Networks* 17:299–318.
———. 1997. "Concurrent Partnerships and the Spread of HIV." *AIDS* 11:641–48.
Morris, M., C. Podhisita, M. Wawer, and M. Handcock. 1996. "Bridge Populations in the Spread of HIV/AIDS in Thailand." *AIDS* 10:1265–71.
Newman, M. E. J. 2002. "Spread of Epidemic Disease on Networks." *Physical Review* E 66 016128.
Newman, M. E. J., S. J. Strogatz, and D. J. Watts. 2001. Random Graphs with Arbitrary Degree Distributions and Their Applications. *Physical Review E* 84:026118.
Phillips, L., J. Potterat, and R. Rothenberg. 1980. "Focused Interviewing in Gonorrhea Control." *American Journal of Public Health* 70:705–8.
Rothenberg, R. B., J. J. Potterat, and D. E. Woodhouse. 1996. "Personal Risk Taking and the Spread of Disease: Beyond Core Groups." *Journal of Infectious Diseases* 174 (suppl. 2): S144–S149.
Rothenberg, R. B., et al. 1997. "Using Social Network and Ethnographic Tools to Evaluate Syphilis Transmission." *Sexually Transmitted Diseases* 25:154–60.
Sattenspiel, L. 1990. "Modeling the Spread of Infectious Disease in Human Populations." *Yearbook of Physical Anthropology* 32:245–76.
Schuster M. A., R. M. Bell, and D. E. Kanouse. 1996. "The Sexual Practices of Adolescent Virgins: Genital Sexual Activities of High School Students Who Have Never Had Vaginal Intercourse." *American Journal of Public Health* 86:1570–76.
Schweizer, T., and D. R. White. 1998. *Kinship, Networks and Exchange*. Cambridge: Cambridge University Press.
Snijders, T. A. B. 1991. "Enumeration and Simulation Methods for 0–1 Matrices with Given Marginals." *Psychometrika* 56:397–417.
St. John, R., and J. Curran. 1978. "Epidemiology of Gonorrhea." *Sexually Transmitted Diseases* 5:81–82.
Turner, C. F., et al. 1998. "Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology." Science 280: 867–73.
Wasserman, S., and K. Faust. 1994. *Social Network Analysis*. Cambridge: Cambridge University Press.
Wasserman, S., and P. Pattison. 1996. "Logit Models and Logistic Regressions for Social Networks: I. An Introduction to Markov Graphs and P*." *Psychometrika* 61:401–25.
Watts, D. 2003. *Six Degrees: The Science of a Connected Age*. New York: W. W. Norton.

This content downloaded from 128.59.161.126 on Fri, 13 Feb 2015 15:01:48 PM
All use subject to JSTOR Terms and Conditions