



HHS Public Access

Author manuscript

Demography. Author manuscript; available in PMC 2016 June 01.

Published in final edited form as:

Demography. 2015 June ; 52(3): 919–942. doi:10.1007/s13524-015-0383-4.

Sexual Mixing in Shanghai: Are Heterosexual Contact Patterns Compatible With an HIV/AIDS Epidemic?

M. Giovanna Merli,

Duke Population Research Institute, Sanford School of Public Policy, Department of Sociology and Duke Global Health Institute, Duke University, Box 90312, Durham, NC 27708, USA
giovanna.merli@duke.edu

James Moody,

Duke Population Research Institute, Department of Sociology, Duke University, Box 90088, Durham, NC 27708, USA

King Abdulaziz University, Saudi Arabia

Joshua Mendelsohn, and

RAND Corporation, 1776 Main Street, Santa Monica, CA 90401, USA

Robin Gauthier

Department of Sociology, Duke University, Box 90088, Durham, NC 27708, USA

Abstract

China's HIV prevalence is low, mainly concentrated among female sex workers (FSWs), their clients, men who have sex with men, and the stable partners of members of these high-risk groups. We evaluate the contribution to the spread of HIV of China's regime of heterosexual relations, of the structure of heterosexual networks, and of the attributes of key population groups with simulations driven by data from a cross-sectional survey of egocentric sexual networks of the general population of Shanghai and from a concurrent respondent-driven sample of FSWs. We find that the heterosexual network generated by our empirically calibrated simulations has low levels of partner change, strong constraints on partner selection by age and education, and a very small connected core, mainly comprising FSWs and their clients and characterized by a fragile transmission structure. This network has a small HIV epidemic potential but is compatible with the transmission of bacterial sexually transmitted infections (STIs), such as syphilis, which are less susceptible to structural breaks in transmission of infection. Our results suggest that policies that force commercial sex underground could have an adverse effect on the spread of HIV and other STIs.

Keywords

Sexual networks; HIV/STI; China; Assortative mating; Network simulation

Introduction

According to the most recent reports on the Chinese HIV/AIDS epidemic, an estimated 780,000 people at the end of 2011 were living with HIV (range: 620,000–940,000) (Ministry of Health 2011, 2012, accounting for 0.058 % of the adult population infected. Among the 48,000 new infections estimated for 2011, heterosexual transmission accounted for 52.2 % of the cases (mainly female sex workers (FSWs), their clients, and the regular partners of clients of sex workers), homosexual transmission accounted for 29.4 %, and the remainder acquired infection mainly through injection drug use (Ministry of Health 2011). The prevailing consensus is that China has joined a cluster of “moderate prevalence” countries, where the epidemic is in slow but steady expansion (Commission on AIDS in Asia 2008a). HIV infections in these countries are mainly driven by FSWs; the fraction of the male population who are clients of FSWs; sex workers’ rate of client turnover; men who have sex with men (MSMs), many of whom marry women and have concurrent sex with men (Ma et al. 2007); and, to a lesser extent, injection drug users (Chin et al. 1998; Commission on AIDS in Asia 2008a, 2008b).

The gains of the Chinese HIV/AIDS epidemic have been attributed to recent, profound changes in sexual attitudes and behaviors that have accompanied China’s rapid pace of social and economic change (Farrer 2002; Sigley and Jeffreys 1999) and to the growth of the commercial sex industry in the past several decades (Lu et al. 2008; Zhang et al. 2008), with a documented reemergence of syphilis and other sexually transmitted infections (STIs) taken as evidence for this transformation in behaviors (Chen et al. 2007; Parish et al. 2003; Tucker et al. 2010). However, if compared with other “moderate-prevalence” Asian countries (such as Indonesia, Malaysia, or Vietnam) with similar or lower levels of male patronage of commercial sex and FSWs client turnover, HIV prevalence in the general population and among sex workers in China is still low (Commission of AIDS in Asia 2008b:38–39). The fraction of FSWs testing positive for HIV in sentinel surveillance sites still does not exceed 0.3 % (Ministry of Health 2012). A recent review of 15 studies of HIV infection among small ($n < 400$) samples of FSWs recruited mainly by convenience sampling revealed a median prevalence rate of 0.6 % and prevalence as high as 10 % only in samples of FSWs who inject drugs (Yan et al. 2011). At present, the recommended approaches for preventing further spread of HIV to the general population include targeting population groups most affected by the epidemic and the organizations able to assist them (Zhang et al. 2008).

Predictions of the spread of HIV in China lay more emphasis on the individual risk behaviors of specific population groups than on the patterns of contact between these groups. Yet, the relationship between HIV infection risks and sexual behavior variables also depends on sexual networks, whereby an individual is linked not only to his/her sexual partners but also to these partners’ partners, facilitating contacts between groups with different levels of sexual activity (Helleringer and Kohler 2007; Morris 1997), and the structure of these networks. Patterns of sexual mixing are especially salient in China because the level of HIV prevalence in the population is low and the probability of a susceptible individual having sexual contacts with an infected partner randomly chosen from the population is very small. In this setting, observable characteristics of partners condition the

probability of HIV infection such that multiple partners chosen from one group may be less risky than a single partner chosen from another group.

The networks that facilitate contacts between population groups and channel susceptible individuals' exposure to infected partners are driven by rules of partner selection. These rules are heavily structured by shared norms as well as social, geographic, and economic factors; and they are defined by observable attributes, such as education, age, race, and regional origin (Laumann et al. 1994). Because, until recently, most sexual activity in China was normatively and socially constrained within marriage, knowledge about rules governing sexual mixing primarily comes from studies of assortative mating in marital partnerships. These studies showed that the status hypergamy by age and socioeconomic status (SES) was prevalent in traditional Chinese society (Thornton and Lin 1994); that later improvements in the status of women relative to men in terms of education and occupational (Hannum and Xie 1994) accompanied by increasing returns to education led to education homogamy in the selection of marital partners (Han 2010); and that starting in the 1990s, worsening women's occupational prospects and postponed marriage for men due to increasing economic pressures brought about a resumption of status attainment through age hypergamy among the post-1990 marriage cohorts, although the probability of crossing more than one five-year age group in the selection of marital partners is very small (Mu and Xie 2011).

China's social and economic transformations of the recent decades have triggered a "sexual revolution" (Farrer 2002; Pan 1994; Sheng 2005; Zha and Geng 1992) characterized by earlier onset of first sex and delays in age at first marriage (Sun 2000); a rise in premarital sex (Gao et al. 2002; Li 2002); a resurgence in the demand and supply of commercial sex (Hershatter 1997; Sigley and Jeffreys 1999); and a surge in extramarital sex behaviors, especially among men coming of age in the 1980s and after (Parish et al. 2007; Tian et al. 2014). Whether the same social, institutional, and economic filters that structure marital ties operate to structure less-committed, less-stable relationships is unknown. Also unknown is the extent to which potentially different rules of selection of nonmarital partners affect patterns of heterosexual mixing and network connectivity in ways that expose more people across different groups to infected partners.

Here, we explore patterns of heterosexual mixing, structural features of sexual networks, and their implications for the progression of HIV over these networks in a Chinese context. To do so, we use unique data on egocentric sexual networks from two surveys of sexual behavior and sexual networks that were concurrently conducted among the general population and among FSWs in Shanghai. We use the observed sexual contact data to simulate large sexual networks that are consistent with observed patterns of mixing by demographic and behavioral categories, numbers of partners, and relationship duration. We then evaluate epidemic potential by introducing a dynamic disease transmission model over these simulated networks to evaluate the proportion of the population exposed to infection.

This work adds to the understanding of patterns of sexual mixing in China and their contribution to the progression of HIV in multiple ways. First, to reflect the growing heterogeneity of heterosexual unions in China, we identify differences in patterns of partner selection across premarital, marital, and extramarital partnerships. Second, we merge the

sociological and epidemiological perspectives to identify key attributes of the social organization of heterosexual partnering—in particular, the sexual mixing of population groups identified by their age, education, and levels of sexual activity, as well as the extent to which these patterns protect or expose susceptible individuals to infected partners. Third, we use a dynamic model of disease transmission that parameterizes behaviors using the temporal sequence, duration of partnerships, and individual risk behaviors grounded in empirical data. Fourth, to improve the utility of a general population survey for modeling the spread of HIV, we rely on additional data collected among FSWs.

We intend the simulations as a heuristic tool to understand whether structural features of heterosexual networks documented in one of China's largest cities are compatible with an HIV epidemic and the conditions required to sustain it. Even if sexual mixing patterns documented in Shanghai cannot represent a country as large and heterogeneous as China, Shanghai is China's second-largest city, at the forefront of changes in sexual norms and behaviors (Farrer 2002; Farrer and Sun 2003). Sexual mixing patterns in Shanghai provide a window on China's regime of sexual relations. Simulations of the sexual network driven by sexual network data collected in one such city can be more informative on the potential for HIV spread in China than arbitrary input parameters chosen to compensate for the dearth of sexual networks data at the national level.

Data

The Shanghai Sexual Networks Survey (SSNS) is a local sexual networks survey conducted between October 2007 and January 2008. The sampling scheme yielded a citywide representative sample of Shanghai 18- to 49-year-old residents with a Shanghai household registration (*hukou*) and migrants (residents without a Shanghai *hukou*). Participation rates in this survey of 56 % for registered residents and 61 % for migrants are not unusual in present-day urban China (de Herr and de Leeuw 2002; Treiman et al. 2009). The total sample size was adjusted for nonresponse. This adjustment yielded 1,192 Shanghai registered residents and 496 migrants. Sampling weights were developed in two stages to compensate for unequal selection probabilities and noncoverage, which involved calibration of the age-sex distribution of the samples of registered residents and migrants to conform to the values of the Shanghai 2005 3 % intercensal sample survey. The adjusted sample is representative of the Shanghai adult population of registered residents and migrants ages 18–49. Additional details on the sampling scheme, adjustment for nonresponse, and construction of weights can be found in the appendix.

An egocentric approach was used to collect data on local sexual networks. Information was collected from respondents on their own and their marital/cohabiting partners' demographic, socioeconomic attributes (e.g., age, marital status, education, previous and present occupation, and income). Information was also solicited from respondents no more than three most-recent nonmarital/noncohabiting partners, including demographic and socioeconomic characteristics of partners, relationship characteristics (e.g., type), duration (start and end dates of partnerships), and behavioral repertoire (frequency of sexual intercourse, type of sexual act, and condom use). The sections of the questionnaire, which cover individual sexual behaviors, marital, and nonmarital sexual histories, were self-

administered to improve item response rates on potentially sensitive topics. All interviewees were administered an informed consent during which they were ensured confidentiality of responses, and then were given a small compensation for successful participation.

To assess the validity of the responses, respondents were asked at the end of each interview whether they would participate in a repeat interview. Respondents who agreed to this request (89 %) provided the frame for the selection of a random subsample of 100 respondents who were administered a repeat interview after a gap of about one month. When we compare two responses on the same respondent, objective items (e.g., age, household registration, education, occupation, and income) have a mean kappa statistic of .93. Less-objective items (e.g., age at first sex, number of entertainment establishments in neighborhood, and travel away from Shanghai) and sensitive items (e.g., marital relationship quality, coital frequency with spouse, and responses on nonmarital partners) both have a mean kappa of .92. Although this is not the best indicator of accurate reporting, those who reported on these items did so consistently.

To describe the sexual behaviors of female sex workers, we rely on data from the Shanghai Study of Women Health (SSWH) conducted concurrently with the SSNS. This study collected information on sexual behaviors and biomarkers for chlamydia and gonorrhea from a sample of 550 FSWs aged 18–45, recruited through respondent-driven sampling (RDS). RDS (Heckathorn 1997) is a chain-referral sampling approach that capitalizes on the network structure of hidden populations to identify and interview subjects and aims to provide a probability-based inferential structure for population representation based on a number of assumptions about the network and the sampling process running over it (Salganik and Heckathorn 2004; Volz and Heckathorn 2008). The advantage of using RDS relative to other sampling methods for hidden populations consists in its ability to reach enhanced coverage of the hard-to-reach pockets of the population of interest, to encourage study participation, and to efficiently and cost-effectively recruit large numbers of survey respondents in a relatively short amount of time (Johnston et al. 2006; Kendall et al. 2008; Qun et al. 2008; Robinson et al. 2006; Weir et al. 2012). However, the validity of RDS estimates rests on an idealized model of how sample respondents make referrals to new respondents and how potential respondents are recruited into the sample. The ability of RDS to generate unbiased estimates of population proportions was shown to be very sensitive to the method's assumptions (Gile and Handcock 2010; McCreesh et al. 2012; Merli et al. 2015; Neely 2009; Yamanis et al. 2013). Some have recommended the use of RDS samples as convenience samples (McCreesh et al. 2012). For the purpose of our simulation analyses, we follow this lead and extract simulation parameters of the FSWs population from naïve sample proportions.

Measures

A total of 840 male and 848 female respondents aged 18–49 were interviewed for the SSNS. A key driver of the network simulations is the number of lifetime sexual partners, also known as “degree” in graph theoretical terms. In the SSNS, this measure of sexual activity was derived from the total number of partners reported in respondents' combined marital and nonmarital¹ partnership histories. Because the total number of partners in the combined

partnership histories was capped at four (no more than three in the nonmarital partnership history plus the current marital partner), the lifetime count of sexual partners asked in a separate question was used as the measure of sexual activity for the 21 respondents who reported the number of lifetime partners in excess of the count of partners in the combined partnership histories. The corresponding measure for FSWs was estimated as the sum of the number of regular clients, onetime clients, and stable partners in the last month as reported in the SSWH.

For our description of patterns of sexual mixing in the Shanghai general population, after excluding 287 respondents who had never had sex and six male respondents who reported homosexual partnerships, we classify respondents and their heterosexual alters in the marital and nonmarital partnership histories by gender on the basis of information contained in the sexual partnership histories. We group partnerships between respondents (ego) and their heterosexual partners (alters) by whether they were premarital (i.e., if they took place prior to the first marriage of the respondent), marital, or extramarital. Extramarital partnerships are those that overlap, by any duration, a marital partnership and are identified from reported start and end dates of nonmarital and marital partnerships. We limit the analyses of marital partnerships to the first marital partnerships of ever-married respondents. This leads to the exclusion of 23 remarriages, 5 nonmarital partnerships following a divorce, and 11 nonmarital partnerships following a second marriage.² This yields a sample of 1,655 unique heterosexual partnerships (ego-alter dyads) formed from the marital and nonmarital sexual histories of 1,389 sexually active male and female SSNS respondents. Of these partnerships, 77.1 % were first-marital, 19.3 % were premarital, and 3.7 % were extramarital. Forty-one percent of men's extramarital partnerships and 20.6 % of premarital partnerships involved an FSW.

Partnerships are classified according to the birth cohort of the respondent and of his/her partner, forming a mixing matrix by five-year birth cohorts. Mixing matrices are also formed by education of respondents and their partner. Both the respondent's education and the education of their partner are recorded into one of four categories: primary, junior high, senior high, and tertiary. The data on all marital dyads provide complete information on education, but data for several premarital and extramarital dyads include only the respondents' education and not the education of their partner. "Don't know" responses on partner education contribute to a distinct category in the mixing matrix and allowed us to assess the extent to which lack of knowledge of educational attainment is a barrier to sexual relations.

¹Nonmarital partners could include previous marital partners of currently remarried respondents or of divorced but currently unmarried respondents.

²The rationale for limiting the analyses to first marriages is because in cross-sectional data on prevailing marriages gathered from multiple birth cohorts, first marriages can lessen biases due to variation in marriage timing and divorce rates, which may affect the degree of resemblance between spouses in a cohort (Kalmijin 1991, 1998; Mare 1991; Qian 1998; Qian and Preston 1993; Raymo and Xie 2000). We do not expect this exclusion to be problematic for our analyses. Because remarriage and divorce are rare in this setting, there is no current evidence of differential risk for STI exposure by marital status. Inclusion of this additional stratifying feature in our models and measures would have created estimation issues as the cell sizes are too small to provide stable mixing estimates. However, because divorce and remarriage are likely to become increasingly prevalent in China, this is an area deserving attention in future work.

Methods

Mixing Matrices

To summarize mixing matrices, we compute a coefficient α , which measures the deviation (or selection bias) of a partnership between men with attribute i and women with attribute j from what we would expect if sexual mixing were random—that is, proportional to population composition. This coefficient is equal to the ratio of observed pairings between men with attribute i and women with attribute j to expected pairings under an independence model in which partners are selected at random given population composition and partnership rates (Morris et al. 2009).

$$\alpha_{ij} = \frac{\log f_{ij}}{\log F_{ij}},$$

where $F_{ij} = M_i + F_{+j} / D_{++}$; f_{ij} is the observed number of partnerships between men with attribute i and women with attribute j ; M_i and F_j are the total numbers of partnerships formed by men of attribute i and women of attribute j ; and D is the total number of partnerships in the sample. Values greater than 1 indicate positive selection bias (relative to random mixing), and values less than 1 indicate negative selection bias (Morris 1991:351).

Simulation Design

To understand the effects of network structure on network connectivity, we generate random networks, using a variant of the randomized mixing algorithm proposed by Newman and Girvan (2003). This algorithm allows us to generate networks with known mixing rates while simultaneously fixing the degree distribution and (within mixing categories) assortative mixing by number of partners (or degree). The algorithm needs two types of information: the mixing probabilities, and the degree distributions by attribute. To construct the network, we start by drawing M edges (ties between two partners or nodes in a network) from the observed mixing distribution by age and education. Each edge is of a mixing type M_{ij} ; hence, $M_j = \sum_i M_{ij}$ is the number of edges involving somebody from category j . Each edge has a “male” and a “female” half. We randomly draw nodes from the attribute-specific degree distribution until the sum of degrees = M_j . We then sort the male and female edge-halves by mixing categories and a random number, and finally merge. This procedure ensures that the mixing distribution of edges matches that of the observed network and maintains the attribute-specific degree distributions. To control assortative mixing by degree, we can include degree in the randomized sort variable (degree + e), with different signs by gender to force disassortativity.³

³The advantages of this randomized draw procedure over other graph simulation approaches—particularly exponential random graph models (ERGM) (e.g., Goodreau et al. 2009)—is mainly computational efficiency. We can generate large networks consistent with our population in mere seconds, which we can then evaluate for connectivity features. Because the information we have is purely at the node and dyad level—including no other edge-dependent features—the resulting models are effectively identical to a dyad-independent ERGM similarly based on node mixing and degree. The addition of the assortativity features creates some level of dependence, but in practice, these are largely determined by birth cohort/education differences in degree.

We start by focusing on network connectivity and the structural properties of the network that influence it. We measure connectivity by the size of the largest connected component, which consists of all nodes that are connected by at least one path over time and the number of paths between these nodes. Connectivity of the network becomes more robust if there are multiple paths around single nodes that could otherwise limit contact through their potential removal. Two paths are node-independent if they have only nodes i and j in common. Multiple node-independent paths represent redundant connections, strengthening the connectivity of the network (Moody and White 2003; Morris et al. 2008). The number of people connected by at least two independent paths—referred to as a *bicomponent*—is the section of the network more strongly connected. We use the largest bicomponent as a measure of the robustness of the potential network core.

Given that we are interested in understanding the global effect of sexual mixing on network connectivity and epidemic potential, we need a network large enough to capture the population-level tendencies in simulation. Because a population-size graph (i.e., billions of nodes) is computationally impossible, we use large networks and explore a range of sizes to identify any trend in the final statistics, which allows us to run the many simulations needed to judge variability. We generate networks of four sizes ranging between $M = 15,000$ and $M = 30,000$ edges. This is the governing size parameter of the network simulation procedure. We are confident that these networks are large enough to prevent any effect of small network sizes in models of component size. Because we do not expect age and education of FSWs to be the most relevant attributes governing their pattern of sexual mixing, we treat FSWs as a distinct mixing category and draw the degree distribution of FSW from the SSWH. Our simulations generate networks that range from 20,000 to 50,000 nodes. Network graphs were produced using PAJEK (De Nooy et al. 2005) and yEd (<http://www.yworks.com/en/index.html>).

The transmission of HIV over a population is a function of the *contact* network, infectivity, and duration of infection. In network simulations that take into account the temporal order of partnerships, setting the value of infectivity in a partnership to 1 generates a reachable path for transmission (e.g., Morris et al. 2009). In this case, the size of the largest connected component captures the maximum number of people who could *potentially* be reached by infection. Because transmission is never perfect, multiple alternate routes available for transmission increase the likelihood of being reached. Combining these features, maximum epidemic potential increases with the size of the largest connected (any reachability) and bicomponent (robust reachability). In our particular case, we rely on the size of the largest component only to gauge the effects of sexual mixing patterns on connectivity. This approach overestimates actual epidemic potential for two reasons. First, in our simulations, reachable nodes in the largest component and bicomponent are defined by cumulative degree distributions informed by number of lifetime partners, with no regard for time. Yet, realistically, one can pass infection only forward—that is, to current, concurrent, or future partners. Second, HIV infection is transmitted with probability per sexual act that is much smaller than 1 (Downs and De Vincenzi 1996) and only for an individual's duration of infection. Over the lifetime of a relationship, this probability can be close to 1 but only in partnerships that are long-lasting and have a high number of sexual acts throughout the

relationship. Transmission probability will be much smaller for one-time sexual encounters, such as those with FSWs.

With these considerations in mind, and to increase the realism of our evaluation of epidemic potential, we simulate a disease transmission model over the simulated network structure. This is a stochastic dynamic network simulation model that uses an investigator-assigned probability for each transmission component (i.e., number of sex acts and transmission probability) and iterates over time. The simulation draws partnerships by duration coded to months from the appropriate mixing matrix cells. Partnership duration and timing match the median of the distributions within each mixing cell of the bundle of original network simulations, and cover approximately 30 years, consistent with the sexual history durations of SSNS respondents. In this way, the networks preserve the population-level cross-sectional statistics (e.g., mean duration of partnership within each mixing cell) of the SSNS, while the formation and dissolution of partnerships is consistent with the sequence and duration of partnerships in the SSNS. HIV is introduced into the simulated network by infecting randomly selected seeds, and new infections are introduced at random throughout the simulation. Seeds are high-degree nodes (mainly FSW). This spontaneous-infection parameter is consistent with the idea of a group of partners of FSWs who are external to the simulated population and who introduce infection into the population at large. This approach generates an average of 8 or 9 extra seeds per simulation setting, with a range of up to 21 extra seeds over the lifetime of each simulation. To estimate the proportion of nodes in the network reached by infection, we model the per-partnership risk of HIV transmission from an infected to a susceptible partner as a function of the infectivity per sex act, the number of exposures per month, and the duration of the relationship.

If the probability of transmission β in a sex act a is a function of the biological transmission risk (β_{ij}) for two members of a discordant couple i and j and behaviors that affect transmission risk, such as the probability that condoms are used within the relationship (c_{ij}), the per-act transmission probability is $\beta_{ij}(1-c_{ij})$. Our simulation works on the micro time scale and uses $\beta_{ij}(1-c_{ij})$ as the transmission probability at each iteration. The total risk of exposure over the lifetime of a partnership (p) would be this probability repeated for the number of sex acts per month (n) over the duration of the relation (T). This would be equal to 1 minus the probability of not transmitting at any event:

$$\beta_p = 1 - (1 - \beta_{ij}(1 - c_{ij}))^{nT}.$$

Table 1 shows the simulation parameters. Disease dynamics are simulated under three values of infectivity per sexual act β_a , ranging from a low value drawn from a strictly controlled study of 525 HIV-discordant European couples (Downs and De Vincenzi 1996) to a high value drawn from studies of individuals with symptomatic AIDS. Partnership characteristics (partnerships can either be with an FSW or with a noncommercial sex partner) determine the behavioral inputs, which are set to vary within each partnership type based on three levels of coital frequency (n) and three corresponding fractions of condom-protected sexual acts (c) with a given type of partner. We also differentiate by the likelihood of infection entering the population from outside the system (two levels), leading to 54

unique parameter settings ($3 \times 3 \times 3 \times 2$). All scenarios preserve the attribute mixing observed in the SSNS data. The outcome at the end of the 30-year simulation period is the proportion of the network population that is infected with HIV.

Results

Observed Data: Mixing Matrices and Degree Distributions

Key drivers of the simulated network structure are patterns of mixing by age and education and degree distributions for each education and birth cohort combination. Figures 1 and 2 show mixing matrices by birth cohort, education, and partnership type. Selection coefficients are presented for each cell of the matrix. Visually, red cells (see the online version of the figures, which appear in color) indicate a positive selection bias, suggesting that a pairing is more likely to happen relative to random mixing; and purple and blue cells indicate a negative selection bias relative to random mixing. Thus, clusters of red cells along the diagonal imply assortative mixing (predominance of within categories ties), clusters of red cells above the diagonal to the right indicate hypergamy (women pairing with partners older or more educated than themselves), and clusters of red cells below the diagonal to the left indicate hypogamy patterns (women pairing with partners younger or less educated than themselves).

For mixing by birth cohort (Fig. 1), premarital sex—a relatively recent phenomenon in China—is prevalent among the youngest cohorts, among whom people mix assortatively with some hypergamous and hypogamous tendencies. Patterns of assortative mixing characterize marital partnerships with some asymmetry because of the propensity of males to pair with somewhat younger female partners. Patterns of extramarital pairings, also a relative recent phenomenon, show both homogamy and some marked propensities of men to choose younger extramarital partners. This represents a marked difference from patterns of marital mixing.

In Fig. 2, premarital patterns of mixing by education demonstrate strong homogamy among the top two educational groups with some activity spreading outward as younger, better-educated cohorts are exposed to similar others in school settings. Patterns of marital mating by education are strongly homogamous, especially at the edges (primary and tertiary education). Extramarital partnerships are quite dispersed with hints of homogamy and hypergamy, as indicated in the online version of the figures by red cells off the right of the diagonal but also some red cells below the diagonal. Education does not matter much in the selection of extramarital partners. About one-half of men's extramarital partnerships in the SSNS involve FSWs, whose educational level is less likely to be known to male respondents.

Figure 3 shows degree distributions by 10-year birth cohorts (column) and education (row) separately for SSNS male and female respondents based on the measure of sexual activity in the SSNS. Among men, degree distributions suggest very dense lower tails, with the majority of respondents reporting one lifetime partner—a feature especially pronounced among older cohorts and men with low educational levels. Distributions extend some in the tail only among younger cohorts and men with higher education, but even these are very

right-skewed, with very few male respondents reporting more than six lifetime partners. For women, the lower tail is denser, with the vast majority reporting one partner, except for some members of the youngest cohorts who report multiple partnerships. The obvious finding from this comparison is that the male degree distribution doesn't balance with the female degree distribution: the number of heterosexual partnerships reported by men exceeds the number reported by women. In this population, about 27 % of partnerships reported by men are unaccounted for in women's reports. This translates into a 1.37 male/female gender ratio of reports. This gender gap in sexual activity reports is consistent with other sexual behavior surveys, which have found that men report up to 60 % more partners than women (Morris 1993; Brewer et al. 2000).

The exclusion of FSWs from the sample is one possible explanation for the imbalance in the male-female degree distributions. In the SSNS, about 8 % of men aged 18–49 report ever paying for sex with a FSW. To compensate for this discrepancy, each man would need four different FSW lifetime partners, which is not unreasonable. Another explanation is a gender gap in reporting with either men overreporting or women underreporting the number of partners, although women would have to underreport by 37 % in order to balance this discrepancy. Another possibility is that the bias is located in the upper tail of the distribution (Morris 1993). Evidence from the SSNS is not inconsistent with this view. Ninety-five percent of men and 99 % of women report fewer than four lifetime partners, and the gender report ratio for this low-degree segment of the population drops to 1.17, suggesting that the imbalance in reports mainly pertains to the upper tail of the distribution.

One way of accounting for the discrepancy in the upper tail of the reporting distribution is to incorporate the degree distribution of FSWs shown in Fig. 4, which is based on the count of clients and stable partners in the last month reported by SSWH respondents. This approach presents one problem, however. Because the degree of FSWs is disproportionately large, and the number of female sex workers is determined by the FSW degree distribution matched to male edges (M_{fsw}) in the mixing matrix, it is unlikely that the FSW degree distribution will fall “within sample.” To address this problem, we test the sensitivity of the simulation results to variation in the FSW degree distribution by using three distributions (max of 30, 60, or 90 partners). Because our simulation model turns on the number of edges M , increasing the degree of FSW lowers their number (because fewer FSWs are needed to have total degree sum to M_{fsw}), and vice versa.

This feature of the model makes it a demand-driven model, where the number of male edges (M_{fsw}) with which to match sex workers drives the size of the FSW population. The number of FSWs as a proportion of the Shanghai female population associated with these three scenarios is slightly higher or lower than 0.5 %, depending on the chosen FSW maximum-degree setting.

Results of Simulations

Simulation of Network Structure—To simulate network structure and its effects on connectivity, we rely on a total of 10,800 simulation runs: 900 runs in each of the three FSW degree settings by four levels of network size ($900 \times 3 \times 4$). Figure 5 shows the distribution of the sizes of the connected components in the simulated networks (left panel) from all

simulation runs. Connected components range in size from less than 1 % of the nodes (typically just isolated dyads) to 18.7 %. More than 80 % of components are very small, and the mean connected component size contains only 0.62 % of the nodes of the network. The largest component ranges from 12.8 % to 18.7 % of the total and averages 16 % irrespective of base network size. Nested within components are bicomponents, whose sizes range from 8 % to 14 % of the largest component and from 1.3 % to 2.1 % of all nodes (results not shown).⁴

The left panel of Fig. 6 shows the largest component of a single run of 22,226 nodes, accounting for 17.5 % of all nodes.⁵ The bicomponent here is 423 nodes, or 10.8 % of the largest component and 1.9 % of all nodes. As shown in the right panel, inspecting one section of the bicomponent revealed by taking an eight-step walk from a randomly chosen node (circled in blue), we observe a string of high-degree nodes, denoted by size and color of node (highlighted in dark green in the online version of the figure) who have contacts with high-degree alters. Taken together, the two panels of Fig. 6 suggest an underlying redundant, albeit small, network core composed by high-degree nodes (probably FSWs) and their high-degree partners (their paying clients).

This core, however, contains long paths that are susceptible to breaks as well as few small cycles (paths that begin and end at the same person, forming robust closed loops that remain connected even if individuals or partnerships are removed). For example, the eight-step walk in the right panel of Fig. 6 shows only two small cycles of eight-step length coming out of the central dark-green node. The extent to which this is a feature of reality, as opposed to a lack of edge-contingent parameters in the simulation, is unknowable. Our intuition is that the geographic clustering and detailed demographic matching, combined with the generally sparse nature of large heterosexual networks, suggests this is not an unreasonable characterization, although we are likely missing small pockets of actors engaged in small clustered communities (e.g., those based on sexual fetish practices), which could create unobserved dense clusters.

Education and age significantly constrain who has high degree and, hence, the nodes in the largest component. Figure 7 shows the proportion of nodes in the largest connected component by degree (x -axis) and demographic characteristics (different lines). Men, FSWs, and people who are better educated or who are younger are more likely to be members of the largest component by virtue of their higher degree. HIV-transmission modeling studies have highlighted the effect of selective (nonrandom) mixing patterns by number of partners on the size and shape of the HIV epidemic (Anderson 1991; Anderson et al. 1990; Gupta et al. 1989; Hyman and Stanley 1988). In the absence of empirical data on assortativity by degree in the SSNS, we arbitrarily let the proportion of partnerships in which both partners have the

⁴We assessed the extent to which variation in the setting of FSW degree affected the sizes of the maximum component and the bicomponent, but their distributions did not vary significantly. The mean of the maximum component varied from 15.7 % to 16.7 % of all nodes with each of the three settings of FSW degree, while the mean of the bicomponent varied from 10.7 % to 9.8 % of the largest component; this suggests that a key feature of these simulated networks is the base frequency distribution of the degree of vertices, not variation in the FSW degree settings. To push this assumption of the model, we ran additional simulations allowing the tail of the FSW degree distribution to reach as high as 250; we found no difference in the epidemic potential measures.

⁵This image is generated by a single run with values chosen to represent the middle of the value ranges used in the complete simulation experiment. We choose a maximum FSW degree of 60 with $M = 27,500$ edges.

same degree vary within each age-education cell between 0 % and 7.5 %. Results (not shown here) show a small negative overall effect of assortativity by degree on component size. The more assortative the mixing is by degree, the smaller the size of the largest connected component, although this relationship varies slightly by FSW degree setting. Overall, differences in component size are quite small, ranging from a predicted mean largest component size of 15.5 % for high-assortativity, low-degree FSW regimes to 17.25 % for high-assortativity, high-degree FSW regimes. The strong association among age, education, and degree constrains the range of potential assortativity by degree and limits its impact on the size of the largest connected component.

Simulation of Disease Transmission Over the Network—Figure 8 shows the distributions of the fraction of network nodes that become infected in scenarios driven by different levels of infectivity per sex act and of monthly coital frequencies and the corresponding proportion of condom-protected sexual acts with noncommercial sex partners and with FSWs. These results indicate that regardless of the behavioral and biological inputs driving the simulations, the network structure constrained by age and education mixing observed in the SSNS significantly limits the spread of HIV infection. The range of the mean of the distribution of the proportion infected generated by the simulations is narrow, between 0.05 % and 0.2 % of all nodes. In scenarios driven by sexual behaviors documented in the SSNS (five sex acts per month with marital partner, and three with FSWs) and a range of infectivity per sex act measured across high- and low-income countries, the fraction of nodes reached by infection ranges between means of 0.06 % and 0.1 %. This fraction does not rise above 0.2 % even when we set the value of infectivity to a very high level of 0.006, consistent with estimates of late-stage infectivity.

Discussion and Conclusions

Our evaluation of HIV epidemic potential in the world's largest population combines considerations of how social and demographic constraints influence the selection of partners and structure patterns of mixing between groups with empirically calibrated network simulations and a disease transmission model simulated over the network. This model mirrors the duration and sequence of observed partnerships in two surveys of sexual behavior and sexual networks concurrently conducted among the general population and FSWs in Shanghai, China. The network structure revealed by the simulations, driven by parameters extracted from the SSNS, mostly consists of small, disconnected components, with a single larger, connected core constrained by age and education mixing with slight modifications by assortativity by degree. This connected set reaches about 16 % of the population and has a reconnected core that never exceeds 2.1 % of the population. This small core, populated by FSWs and their highly sexually active male partners, is characterized by long paths and few short cycles through which transmission may occur, highlighting the structural fragility of the network especially for a disease such as HIV, which is not highly infectious (Downs and De Vincenzi 1996). HIV spread simulated over this network never reaches more than 0.2 % of all nodes, even under a scenario of high infectivity most favorable to HIV transmission.

This network structure bears little resemblance to a structure compatible with broad HIV diffusion, such as, the large, robust network structure characterized by many short cycles documented by the Likoma Network Study on a small island on Lake Malawi, where 10.6 % of female respondents and 4.7 % of male respondents tested positive for HIV (Helleringer and Kohler 2007). Moving to an Asian context, a regime of sexual relations compatible with large HIV potential would be one in which demand for commercial sex and FSWs' client turnover are high (i.e., FSWs sell sex to a large number of distinct men), and men have non-exclusive sexual relations with multiple FSWs. In this regime, infected FSWs would transmit HIV to their clients, who in turn would transmit the virus back into the small but increasingly infected pool of sex workers. Such a regime would be consistent with a number of short cycles in the reconnected core making transmission easier, similar to the one that characterized the sexual network structure of HIV epidemics in Thailand or Cambodia. In fact, in those countries, national HIV prevalence would have reached levels as high as 8 % to 10 % had there been no increase in condom use and no reduction in levels of patronage of commercial sex (Commission of AIDS in Asia 2008b). In the SSNS, the proportion of men who buy sex and client turnover for FSWs are simply not high enough to generate the network structure required to sustain an HIV/AIDS epidemic.

Our conclusions regarding the restraining effect of the heterosexual network structure documented in Shanghai on the spread of HIV are valid despite the simplifying assumptions of our models regarding the structure of the network and disease transmission running over it. First, a more precise accounting of unobserved factors—which, in the Chinese context, would further constrain patterns of sexual mixing between groups in addition to age and education (e.g., urban/rural residence)—would only reduce the size of the largest connected component.

Second, even though our stylized disease transmission model cannot account for the natural history of HIV, it is unlikely to underestimate the proportion reached by infection. The natural history of HIV is such that HIV infectivity spikes during the first few months after the onset of infection, drops thereafter to very low levels during a latent long stage, only to rise again with the onset of symptomatic AIDS because of high viral loads observed for the early and late periods of infection (e.g., Pilcher et al. 2004; Wawer et al. 2005). Our worst-case scenario infectivity value of 0.006 drawn from studies of infectivity during the late symptomatic stage is three times the average transmission probability per unprotected sexual act of 0.002 during early acute infection implied by data for U.S. couples (Pilcher et al. 2004; Pinkerton 2008) and five times the average level of HIV infectivity of 0.0012 estimated for European couples during the latent stage (Downs and De Vincenzi 1996). A less-stylized disease transmission model that accounts for the sensitivity of the transmission probability to changes in the amount of virus shed over duration of infection would not have made it more difficult for us to conclude that current patterns of sexual mixing are not compatible with an HIV epidemic.

Third, our simulation model is driven by cumulative degree distributions informed by the number of lifetime partners in the general population and matching males who report ties to a population of FSWs with degree distributions based on the number of partners in the last month. Our approach is pragmatic. We populated our simulation with enough FSWs to meet

the observed general population demand. We did not scale up the degree of FSW to the last-year number of partners for three reasons. First, the number of sex partners in the last month was the only measure of degree collected in the SSWH, so we have no accurate measure over longer time horizons. Second, field reports suggest that much of the FSW volume in any given city is driven by out-of-town visitors, which would be represented as outside our simulation boundary. We accounted for this sort of boundary permeability using a spontaneous infection parameter, allowing disease to enter the population without a known seed. Third, because our model is demand-driven, inflating the FSW degree distribution (by multiplying by 12 months) would only have produced the effect of decreasing the number of FSWs (each connecting to more men). Although this could have a “shortcut” effect on infection potential, it is difficult to know what an accurate degree distribution of FSWs would be. For these reasons, we implemented a series of sensitivity runs, raising FSW degree to more than 250 (compared with 90 in the highest main simulations). This had no effect on component size, redundancy, or simulated infection spread because, we think, the vast majority of FSW-related edges already fall within the most connected regions of the network. As Fig. 7 demonstrates, degree is the best predictor of being in the largest connected component, and FSWs are always the highest degree nodes in the simulation.

More relevant to the epidemiology of HIV in the Chinese context, our model cannot account for the role that injection drug use might play in fueling the heterosexual spread of infection. Injection drug use is a highly efficient mode of transmission, which may increase the size of the epidemic independent of heterosexual transmission or raise the probability of HIV transmission in a partnership through the interaction between sexual transmission and needle sharing. The incorporation of the auxiliary role of injection drug use in heterosexual networks would have required rare data on needle-sharing behaviors and networks and greater sophistication of our modeling strategy. Accounting for injection drug use would have probably increased the simulated proportion of infected nodes.

Finally, the prevalence of premarital and extramarital partnerships, for which age and education constraints may matter less, and the prevalence of divorce and remarriage are still relatively low in China. Because our models do not estimate a future trajectory for an increase in premarital, remarriage, or extramarital relations, they cannot directly speak to future trends in the contact structure. However, a future rise in these types of partnerships with new cohorts replacing older more conservative ones might increase the rate of partner change and the fluidity of mixing groups giving the epidemic greater reach into the general population. Similarly uncertain are the effects of dynamic social and economic changes on the demand and supply of commercial sex, although these may well lead China’s epidemic in the future to derive its momentum from significant levels of HIV transmission during unsafe paid sex. Future detailed empirical work is needed to examine the trends in these potentially higher-risk populations.

Our conclusions may be different for bacterial STIs with different natural histories than HIV. The network structure consistent with patterns of sexual mixing documented in the SSNS data, with a small core populated by FSWs and their clients but with long paths susceptible to breaks in transmission and few short cycles, might be more compatible with the spread of syphilis. Because of its much higher transmission probability per sexual act—

Mitchell et al. (2013) put the transmission probability of syphilis at 0.3 based on prospective controlled trials reviewed by Garnett et al. (1997)— syphilis transmission is less susceptible to structural breaks than HIV. Indeed, recent studies conducted in China have reported high syphilis prevalence among FSWs and their clients as well as in the general population (Chen et al. 2007, 2012; Tucker et al. 2010).

Attention should be paid to members of the core high-degree individuals, including FSWs with high client turnover and patrons of FSWs in the general population, largely because commercial sex networks are a likely entryway through which exogenous sources of infection could enter the network. Policies that force commercial sex underground could have an adverse effect on the spread of HIV because, barring a drop in demand, a smaller number of FSWs with high degree to meet demand will create shortcuts in transmission and more robust transmission routes. At the same time, any widespread increase in the prevalence of extramarital relations will make a generalized connected core much larger. These changes are not unrealistic and combine to push strongly toward policies focused on minimizing concurrent relations because this creates more potential transmission routes.

Acknowledgments

Collection of the survey data analyzed here was undertaken by the first author in collaboration with Ersheng Gao and Xiaowen Tu at Fudan University School of Public Health, and Anan Shen at the Shanghai Academy of Social Sciences. The Shanghai Sexual Network Survey (SSNS) was funded by NICHD/NIDA Grant R21HD047521, supplemented by two smaller grants from NICHD (Merli, PI). The Shanghai Study of Women Health (SSWH) was funded by a Ford Foundation grant to Ersheng Gao. The SSNS sample was designed in consultation with William Kalsbeek. Data analyses were funded in part by NICHD Grants R01 HD068523 (Merli, PI) and R01 HD075712 (Moody, PI). We thank the seminar participants of the Center for Studies in Demography and Ecology at the University of Washington and of the Population Research Center at the University of Texas at Austin, Ashton Verdery, and two anonymous reviewers for constructive comments.

Appendix

Sampling Scheme, Adjustment For Nonresponse and Development of Sampling Weights

Sampling Scheme

The samples of Shanghai registered residents and migrants were selected as random subsamples of Shanghai registered residents and migrants from a stratified multistage clusters sample screened by the Shanghai Statistical Bureau for the 2005 3 % intercensal survey of the Shanghai population.

The design of the 3 % sample produced a stratified two-stage sample of persons, where 963 neighborhood committees (NCs) consisting of 2,000–5,000 persons were selected as primary sampling units (PSUs), with probabilities proportional to their estimated number of residents and migrants. Explicit stratification of NCs was done by the 19 districts/counties (18 urban districts and 1 rural county). NC selection in the first stage was also implicitly stratified within each district/county by the estimated ratio of resident population to migrant population in sorting the NCs for probability proportional to size (PPS) systematic selection within each district/county. Allocation of the 963 NCs among districts/county was proportionate to the population of the stratified subsample of NCs. In the second sampling

stage, all persons (and households) were included within 2,151 small groups (SGs, approximately 100 persons each) that were chosen within each sample NC by unstratified simple random sampling.

The subsamples of residents and migrants for the SSNS was selected from the 3 % sample to yield a stratified four-stage sample of Shanghai residents and migrants aged 18–49. In Stage 1, for both subsamples, 50 NCs were first subsampled from the 963 NCs selected for the 3 % sample. This was accomplished by simple random sampling without replacement within each of the following three groups of the 19 districts/counties used for the 3 % sample: central city, inner suburbs, and outer suburbs. Allocation in the stratified subsample of NCs was proportionate. In Stage 2, exactly two of the two or three SGs that were selected for the second stage of the 3 % sample were retained for the SSNS subsample, yielding 100 SGs subsampled. Separate subsamples of 18- to 49-year-old residents and migrants were then subsampled in the two remaining subsampling stages. In Stage 3, for the resident subsample, 12 registered households were recruited within each selected SG using a currently updated list of household addresses in the SG as the sampling frame. In Stage 4, one 18- to 49-year-old household resident was randomly chosen from among those living in each participating household using a conventional Kish table. For the migrant subsample, a similar procedure was used with five migrant households recruited per SG using a currently updated list of household addresses with at least one migrant present.

Adjustment for Nonresponse

Among the 1,200 Shanghai registered residents and 500 migrants identified in the samples, participation rates were 56 % for Shanghai registered residents and 61 % for migrants; 17.7 % Shanghai registered residents and 17.8 % migrants refused to be interviewed, 14 % and 12 % did not participate because of failure to reach them, and 3.1 % and 3 % did not participate for other reasons. No reason was provided for nonparticipation for the remaining 9.7 % and 5.2 % of the samples. To prevent unit nonresponse from affecting the size of the samples of Shanghai residents and migrants, field substitution (Chapman 1983) was used to compensate for unit nonresponse. In the third sampling stage, within each SG, assigned and reserved samples of potential respondent households were selected according to random procedures. Respondents drawn from the assigned samples who did not participate in the survey were replaced with respondents drawn from reserve samples. Because substitution can easily be abused and the statistical integrity of the final sample can be compromised, we developed and rigorously enforced strict rules for replacing assigned addresses (for example, establishing when to call, what to say and do, and how many attempts to make before substitution). These were emphasized to both interviewers and supervisors at field training. A comparison of basic demographic characteristics of initial nonrespondents with those of respondents drawn from reserve samples and of participating respondents initially assigned for recruitment revealed no significant difference with respect to age and gender. This adjustment procedure yielded a total sample size adjusted for nonresponse of 1,192 Shanghai registered residents and 496 migrants.

Construction of Sampling Weights

Sampling weights were developed in two stages. The first stage of weighting consisted of computing the overall selection probability for each respondent in either the resident or migrant sample as the product of the stage-specific selection probabilities computed for each sampling stage (NC, SG, household, and household member). The individual sampling weight— $w_{respondent}$ —was computed as the inverse of the overall probability of selection. The second stage of weighting involved a further adjustment to this weight to make the resulting weighted estimates from the sample conform to known population values by age and sex for registered residents and migrants in the 2005 3 % intercensal survey. This adjustment was done for the samples of registered residents and migrants as follows: (a) Determine the unweighted sample frequency for each cell n_{cell} defined by the cross-classified calibration variables (age and sex); (b) Determine the population count of 18- to 49-year-olds in the 3 % intercensal survey for each cell N_{cell} defined in the previous step; (c) Compute the calibration adjustment for all respondents in a cell as

$$Adjustment_{cell} = \frac{N_{cell}}{\text{sum of } W_{respondent} \text{ over all } n_{cell} \text{ respondents}};$$

and (d) Compute the final calibrated weight for any respondent in a cell as

$$W_{respondent,calibrated,cell} = (W_{respondent,cell})(Adjustment_{cell}).$$

References

- Anderson, RM. Mathematical models of the potential demographic impact of AIDS in Africa. In: Piot, P.; Kapita, BM.; Were, JBO., editors. AIDS in Africa. London, UK: Current Science; 1991. p. 37-44.
- Anderson R, Gupta S, Ng W. The significance of sexual partner contact networks for the transmission dynamics of HIV. *Journal of Acquired Immune Deficiency Syndromes*. 1990; 3:417–429. [PubMed: 2179528]
- Brewer DD, Potterat JJ, Garrett SB, Muth SQ, Roberts JM, Kaprzyk D, Darrow WW. Prostitution and the sex discrepancy in reported number of sexual partners. *Proceedings of the National Academy of Science*. 2000; 97:12385–12388.
- Chapman, D. The impact of substitution on survey estimates in incomplete data in sample surveys. In: Madow, W.; Olkin, I.; Rubin, D., editors. *Theory and bibliographies*. Vol. II. New York, NY: National Academy of Sciences, Academic Press; 1983. p. 45-61.
- Chen XS, Wang QQ, Yin YP, Liang GJ, Jiang N, Yang LG, Wang BX. Prevalence of syphilis infection in different tiers of female sex workers in China: Implications for surveillance and interventions. *BMC Infectious Diseases*. 2012; 12(84):1–7. [PubMed: 22214291]
- Chen Z, Zhang G, Gong X, Lin C, Fao X, Liang G, Cohen MS. Syphilis in China: Results of a national surveillance programme. *Lancet*. 2007; 369:132–369. [PubMed: 17223476]
- Chin J, Bennett A, Mills S. Primary determinants of HIV prevalence of Asian-Pacific countries. *AIDS*. 1998; 12(Suppl. B):S87–S91. [PubMed: 9679633]
- Commission on AIDS in Asia. *Redefining AIDS in Asia: Technical annex*. UNAIDS Bangkok, Thailand: Regional Support Team for Asia and the Pacific; 2008a.
- Commission on AIDS in Asia. *Redefining AIDS in Asia: Crafting an effective response*. Oxford, UK: Oxford University Press; 2008b.

- de Heer, W.; de Leeuw, E. Trends in household survey nonresponse: A longitudinal and international comparison. In: Groves, RM.; Dillman, DA.; Eltinge, JL.; Little, RRJ., editors. Survey nonresponse. New York, NY: Wiley; 2002. p. 41-54.
- De Nooy, W.; Mrvar, A.; Batagelj, V. Exploratory social network analysis with Pajek. Cambridge, UK: Cambridge University Press; 2005.
- Downs AM, De Vincenzi. Probability of heterosexual transmission of HIV: Relationship to the number of unprotected sexual contacts. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*. 1996; 11:388–395. [PubMed: 8601226]
- Farrer, J. Opening up: Youth sex culture and market reform in Shanghai. Chicago, IL: University of Chicago Press; 2002.
- Farrer J, Sun Z. Extramarital love in Shanghai. *The China Journal*. 2003; 50:1–36.
- Gao, E.; Lou, C.; Tu, X. Reproductive health of unmarried Chinese youth. In: Gao, E.; Lou, C.; Tu, X.; Shah, IH., editors. The reproductive health of adolescent and unmarried youth: Current status, future trends and strategies. Shanghai, China: The Second Military Medical School. (In Chinese); 2002. p. 16-26.
- Garnett GP, Aral SO, Hoyle DV, Cates W, Anderson RM. The natural history of syphilis. Implications for the transmission dynamics and control of infection. *Sexual Transmitted Diseases*. 1997; 24:185–200.
- Gile KJ, Handcock MS. Respondent-driven sampling: An assessment of current methodology. *Sociological Methodology*. 2010; 40:285–327. [PubMed: 22969167]
- Goodreau SM, Kitts JA, Morris M. Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography*. 2009; 46:103–125. [PubMed: 19348111]
- Gupta S, Anderson R, May R. Networks of sexual contacts: Implications for the pattern of spread of HIV. *AIDS*. 1989; 3:807–817. [PubMed: 2517202]
- Han H. Trends in educational assortative marriage in China from 1970 to 2000. *Demographic Research*. 2010; 22:733–770. (article 24).
- Hannum E, Xie Y. Trends in educational gender inequality in China: 1949–1985. *Research in Social Stratification and Mobility*. 1994; 13:73–98.
- Heckathorn DD. Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*. 1997; 44:174–199.
- Helleringer S, Kohler HP. Sexual network structure and the spread of HIV in Africa: Evidence from Likoma Island, Malawi. *AIDS*. 2007; 21:2323–2332. [PubMed: 18090281]
- Hershatter, G. Dangerous pleasures: Prostitution and modernity in twentieth-century Shanghai. Berkeley: University of California Press; 1997.
- Hyman JM, Stanley EA. Using mathematical models to understand the AIDS epidemic. *Mathematical Biosciences*. 1988; 90(1/2):415–473.
- Johnston LG, Sabin K, Mai TH, Pham TH. Assessment of respondent driven sampling for recruiting female sex workers in two Vietnamese cities: Reaching the unseen sex worker. *Journal of Urban Health*. 2006; 83(Suppl.):6–28.
- Kalmijn M. Trends in religious and educational homogamy. *American Sociological Review*. 1991; 56:786–800.
- Kalmijn M. Inter-marriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology*. 1998; 24:395–421.
- Kaufman J, Jing J. China and AIDS: The time to act is now. *Science*. 2002; 296:2339–2340. [PubMed: 12089428]
- Kendall C, Kerr LR, Gondim RC, Werneck GL, Macena RH, Pontes MK, McFarland W. An empirical comparison of respondent driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil. *AIDS and Behavior*. 2008; 12:97–104.
- Laumann, EO.; Gagnon, JH.; Michael, RT.; Michaels, S. The social organization of sexuality: Sexual practices in the United States. Chicago, IL: University of Chicago Press; 1994.

- Li, A. Sexual attitudes and behaviors of university students in Beijing, China. In: Gao, E.; Lou, C.; Tu, X.; Shah, IH., editors. *The reproductive health of adolescent and unmarried youth: Current status, future trends and strategies*. Shanghai, China: The Second Military Medical School; 2002. p. 27-39. (In Chinese)
- Lu L, Jia M, Ma Y, Yang L, Chen Z, Ho DD, Zhang L. The changing face of HIV in China. *Nature*. 2008; 455:609–611. [PubMed: 18833270]
- Ma X, Zhang Q, He X, Sun W, Yue H, Chen S, McFarland W. Trends in prevalence of HIV, syphilis, hepatitis C, hepatitis B, and sexual risk behavior among men who have sex with men. Results of 3 consecutive respondent-driven sampling surveys in Beijing, 2004 through 2006. *Journal of Acquired Immune Deficiency Syndrome*. 2007; 45:581–587.
- Mare RD. Five decades of educational assortative mating. *American Sociological Review*. 1991; 56:15–32.
- McCreesh N, Frost SD, Seeley J, Katongole J, Tarsh MN, Ndunguse R, White RG. Evaluation of respondent-driven sampling. *Epidemiology*. 2012; 23:138–147. [PubMed: 22157309]
- Merli MG, Moody J, Smith J, Li J, Weir S, Chen X. Challenges to recruiting representative samples of female sex workers in China using respondent driven sampling. *Social Science & Medicine*. 2015; 125:79–93. [PubMed: 24834869]
- Ministry of Health, People's Republic of China, Joint United Nations Programme on HIV/AIDS, World Health Organization. 2011 estimates for the HIV/AIDS epidemic in China. Beijing, China: National Center for AIDS/STD Control and Prevention, China CDC; 2011.
- Ministry of Health, People's Republic of China. 2012 China AIDS response progress report. 2012. Retrieved from http://www.unaids.org/sites/default/files/en/dataanalysis/knownyourresponse/countryprogressreports/2012countries/ce_CN_Narrative_Report%5B1%5D.pdf
- Mitchell KM, Cox AP, Mabey D, Tucker JD, Peeling RW, Vickerman P. The Impact of syphilis screening among female sex workers in China: A modelling study. *PLOS One*. 2013; 8(1):e55622. [PubMed: 23383249]
- Moody J, White DR. Social cohesion and embeddedness. *American Sociological Review*. 2003; 68:103–127.
- Morris M. A log-linear modeling framework for selective mixing. *Mathematical Biosciences*. 1991; 107:349–377. [PubMed: 1806123]
- Morris M. Telling tails explain the discrepancy in sexual partner reports. *Nature*. 1993; 365:437–440. [PubMed: 8413586]
- Morris M. Sexual networks and HIV. *AIDS 97: Year in Review*. 1997; 11(Suppl. A):S209–S216.
- Morris, M.; Goodreau, S.; Moody, J. Sexual networks, concurrency, and STD/HIV. In: Holmes, KK.; Sparling, PF.; Stamm, WE.; Piot, P.; Wasserheit, JN.; Corey, L.; Cohen, MS.; Watts, DH., editors. *Sexually Transmitted Diseases*. 4th Ed.. New York, NY: McGraw-Hill; 2008. p. 109-126.
- Morris M, Kurth AE, Hamilton DT, Moody J, Wakefield S. Concurrent partnerships and HIV prevalence disparities by race: Linking science and public health practice. *American Journal of Public Health*. 2009; 99:1023–1031. [PubMed: 19372508]
- Mu, Z.; Xie, Y. *Marital homogamy in China: A reversal in trends in the reform era?* (Report 11-742). Ann Arbor, MI: Population Studies Center, University of Michigan, Institute for Social Research; 2011.
- National Intelligence Council. *The next wave of HIV/AIDS: Nigeria, Ethiopia, Russia, India and China* (ICA 2002-04 D). Washington, DC: NIC; 2002. Retrieved from <http://fas.org/irp/nic/hiv-aids.html>
- Neely, WW. *Statistical theory for respondent-driven sampling* (Unpublished doctoral dissertation). University of Wisconsin–Madison; 2009.
- Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Physical Review E*. 2003; 69(2):1–16.
- Pan S. A sex revolution in current China. *Journal of Psychology & Human Sexuality*. 1994; 6(2):1–14.
- Parish W, Laumann EO, Cohen MS, Pan S, Zheng H, Hoffman I, Ng KH. Population-based study of chlamydial infection in China: A hidden epidemic. *Journal of the American Medical Association*. 2003; 289:1265–1273. [PubMed: 12633188]
- Parish W, Laumann EO, Mojola SA. Sexual behavior in China: Trends and comparisons. *Population and Development Review*. 2007; 33:729–756.

- Pilcher CD, Tien HC, Eron JJ Jr, Vemazza PL, Leu SY, Stewart PW, Cohen MS. Brief but efficient: Acute HIV infection and the sexual transmission of HIV. *Journal of Infectious Diseases*. 2004; 189:1785–1792. [PubMed: 15122514]
- Pinkerton SD. Probability of HIV transmission during acute infection in Rakai, Uganda. *AIDS and Behavior*. 2008; 12:677–684. [PubMed: 18064559]
- Qian Z. Changes in assortative mating: The impact of age and education, 1970–1990. *Demography*. 1998; 35:279–292. [PubMed: 9749320]
- Qian Z, Preston SH. Changes in American marriage, 1972 to 1987: Availability and forces of attraction by age and education. *American Sociological Review*. 1993; 58:482–495.
- Qun H, Wang Y, Li Y, Zhang Y, Lin P, Yang F, McFarland W. Accessing men who have sex with men through long-chain referral recruitment, Guangzhou, China. *AIDS and Behavior*. 2008; 12(4 Suppl.):S93–S96. [PubMed: 18389358]
- Raymo JM, Xie Y. Temporal and regional variation in the strength of educational homogamy. *American Sociological Review*. 2000; 65:773–781.
- Robinson WT, Risser JMH, McGoy S, Becker AB, Rehman H, Jefferson M, Tortu S. Recruiting injection drug users: A three site comparison of results and experiences with respondent driven and target sampling procedures. *Journal of Urban Health*. 2006; 83:29–38.
- Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations: Using respondent-driven sampling. *Sociological Methodology*. 2004; 24:193–239.
- Sheng, X. Chinese families. In: Adams, BN.; Trost, J., editors. *Handbook of world families*. Thousand Oaks, CA: Sage; 2005. p. 99-128.
- Sigley G, Jeffreys E. On “sex” and “sexuality” in China: A conversation with Pan Suiming. *Bulletin of Concerned Asian Scholars*. 1999; 31(1):50–58. [PubMed: 22279648]
- Sun, J. Case study, China, People’s Republic of: communication and advocacy strategies adolescent reproductive and sexual health. Bangkok: UNESCO PROAP; 2000. <http://unesdoc.unesco.org/images/0012/001246/124698e.pdf>
- Thornton, A.; Lin, H. *Social change and the family in Taiwan*. Chicago, IL: University of Chicago Press; 1994.
- Tian FF, Merli MG, Qian Z. Job mobility and extramarital sex in reform-era urban China: Evidence from Shanghai. *Chinese Sociological Review*. 2014; 46(1):60–82.
- Treiman, DJ.; Hu, PP.; Lu, Y.; Mason, WM.; Qi, Y.; Song, S. Determinants and consequences of peasant labor migration in contemporary China; Beijing, China. Paper presented at the Spring Meeting of the International Sociological Association Research Committee on Social Stratification and Mobility; 2009 May.
- Tucker JD, Chen XS, Peeling RW. Syphilis and social upheaval in China. *New England Journal of Medicine*. 2010; 362:1658–1661. [PubMed: 20445179]
- Volz E, Heckathorn DD. Probability-based estimation theory for respondent driven sampling. *Journal of Official Statistics*. 2008; 24:79–97.
- Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, Quinn TC. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *Sexually Transmitted Diseases*. 2005; 191:1403–1409.
- Weir S, Merli MG, Li J, Gandhi AD, Neely WW, Edwards JK, Chen X. A comparison of respondent-driven and venue-based sampling of female sex workers in Liuzhou, China. *Sexually Transmitted Infections*. 2012; 88(Suppl. 2):i95–i101. [PubMed: 23172350]
- Yamanis N, Merli MG, Neely WW, Tian FF, Moody J, Tu X, Gao E. An empirical analysis of the impact of recruitment patterns on RDS estimates among a socially ordered population of female sex workers. *Sociological Methods and Research*. 2013; 42:392–425.
- Yan H, Poon AN, Zhang C. HIV/STI prevention interventions targeting FSWs in China: A systematic literature review. *AIDS Care*. 2011; 23(Suppl.1):54–65. [PubMed: 21660751]
- Zha B, Geng W. Sexuality in urban China. *Australian Journal of Chinese Affairs*. 1992; 28:1–20.
- Zhang K, Detels R, Liao S, Cohen M, Yu D. China’s HIV/AIDS epidemic: Continuing challenges. *Lancet*. 2008; 372:1791–1793. [PubMed: 18930521]
- Zhang K, Ma S. Epidemiology of HIV in China. *BMJ*. 2002; 324:803–804. [PubMed: 11934762]

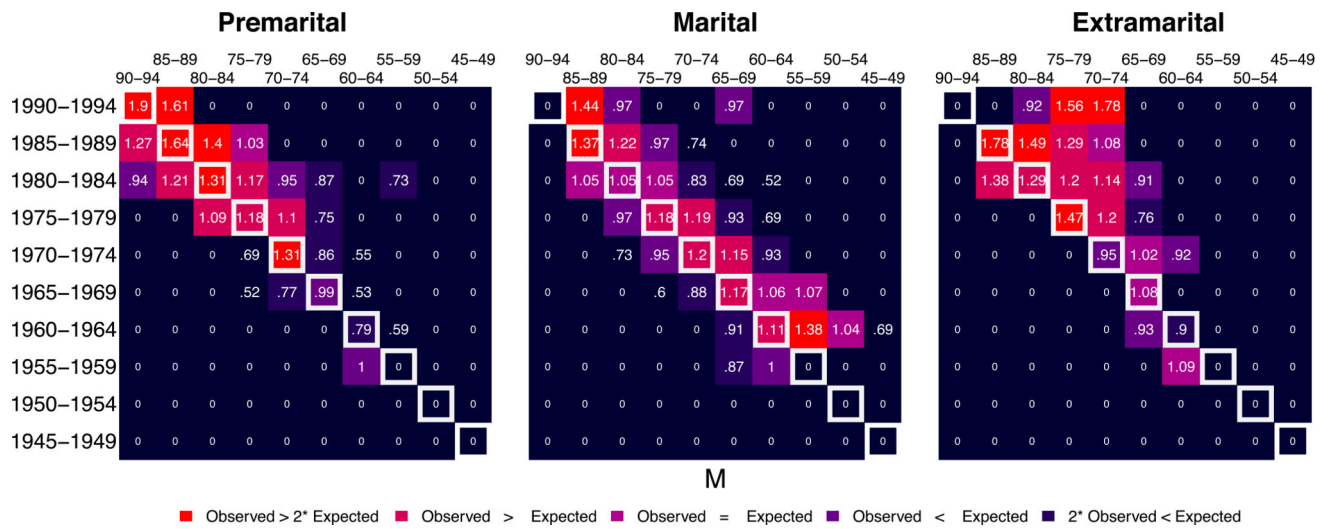
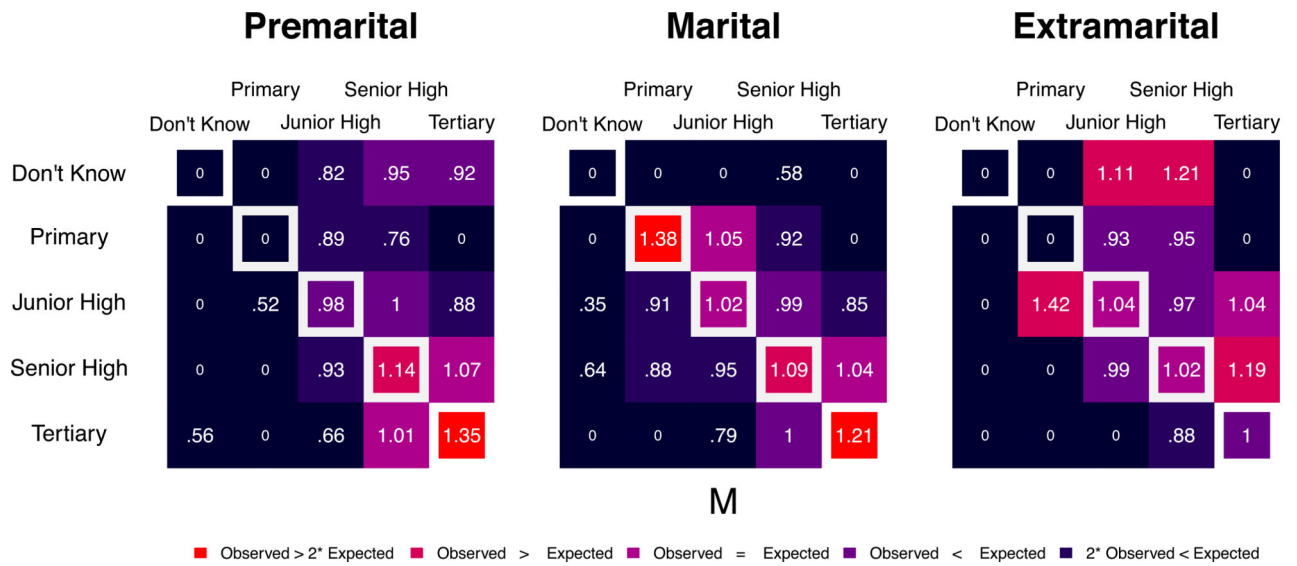


Fig. 1.
Heterosexual mixing by birth cohort in the SSNS



F

Fig. 2.
Heterosexual mixing by education in the SSNS

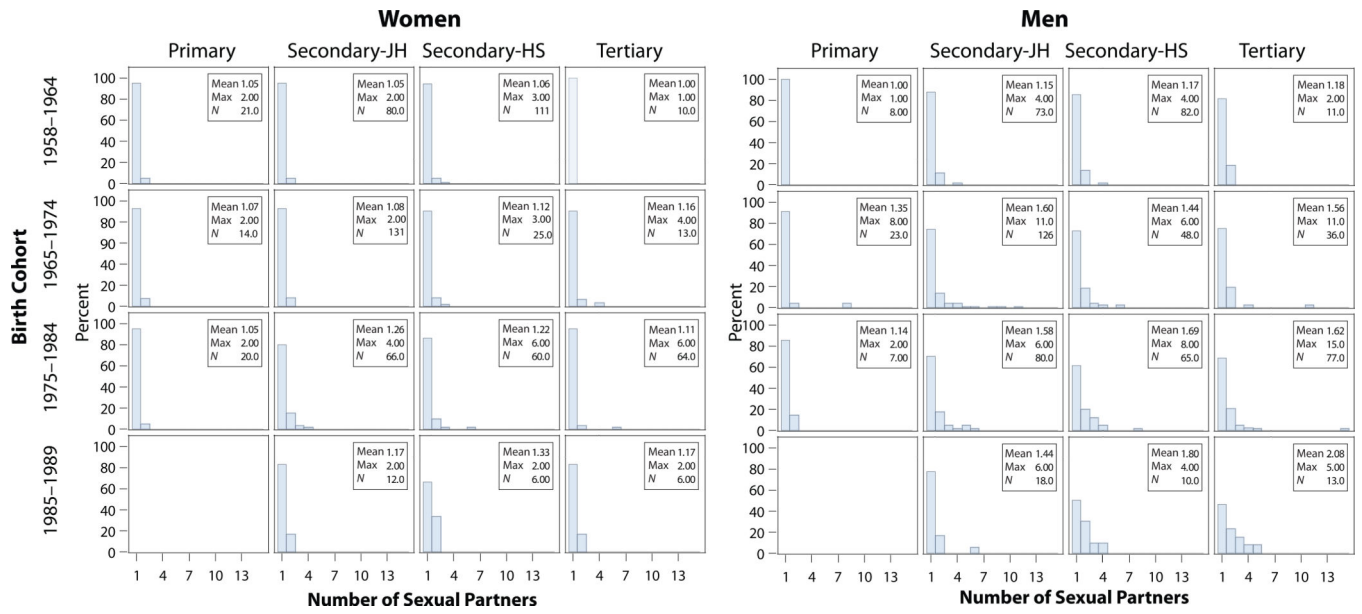


Fig. 3. Observed degree distribution, by birth cohort and education, Shanghai men and women.
Source: SSNS

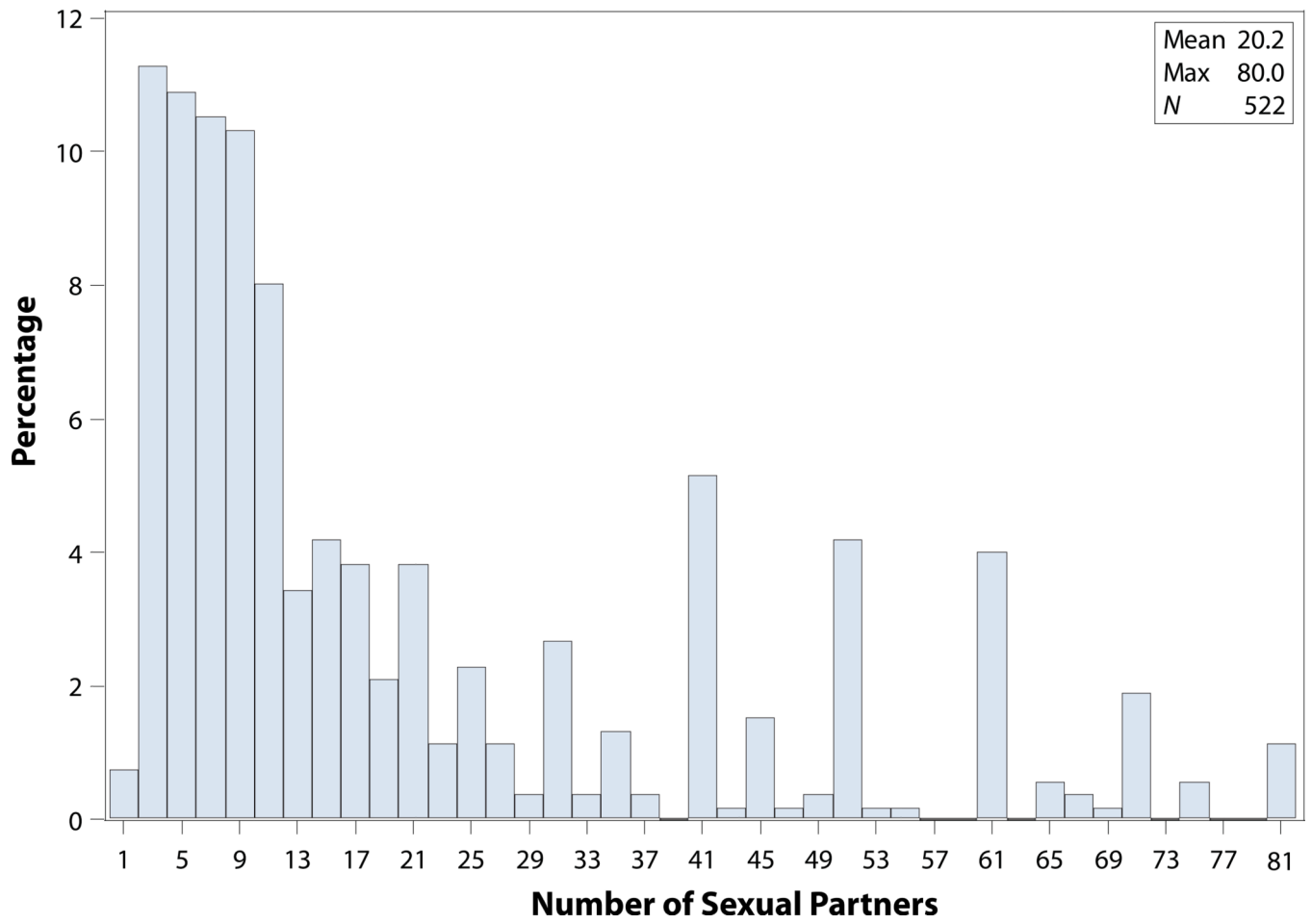


Fig. 4. Observed degree distribution of female sex workers, Shanghai. *Source:*

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

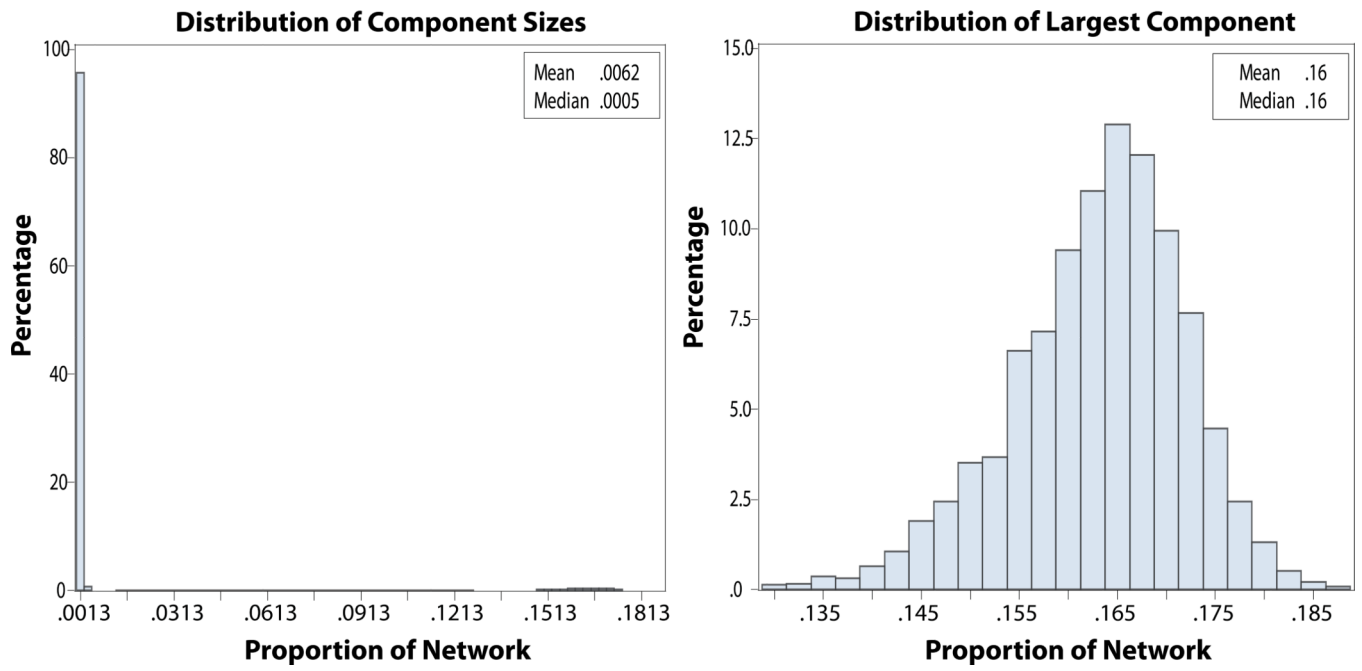


Fig. 5. Size distribution of connected components and of largest component as proportions of all nodes

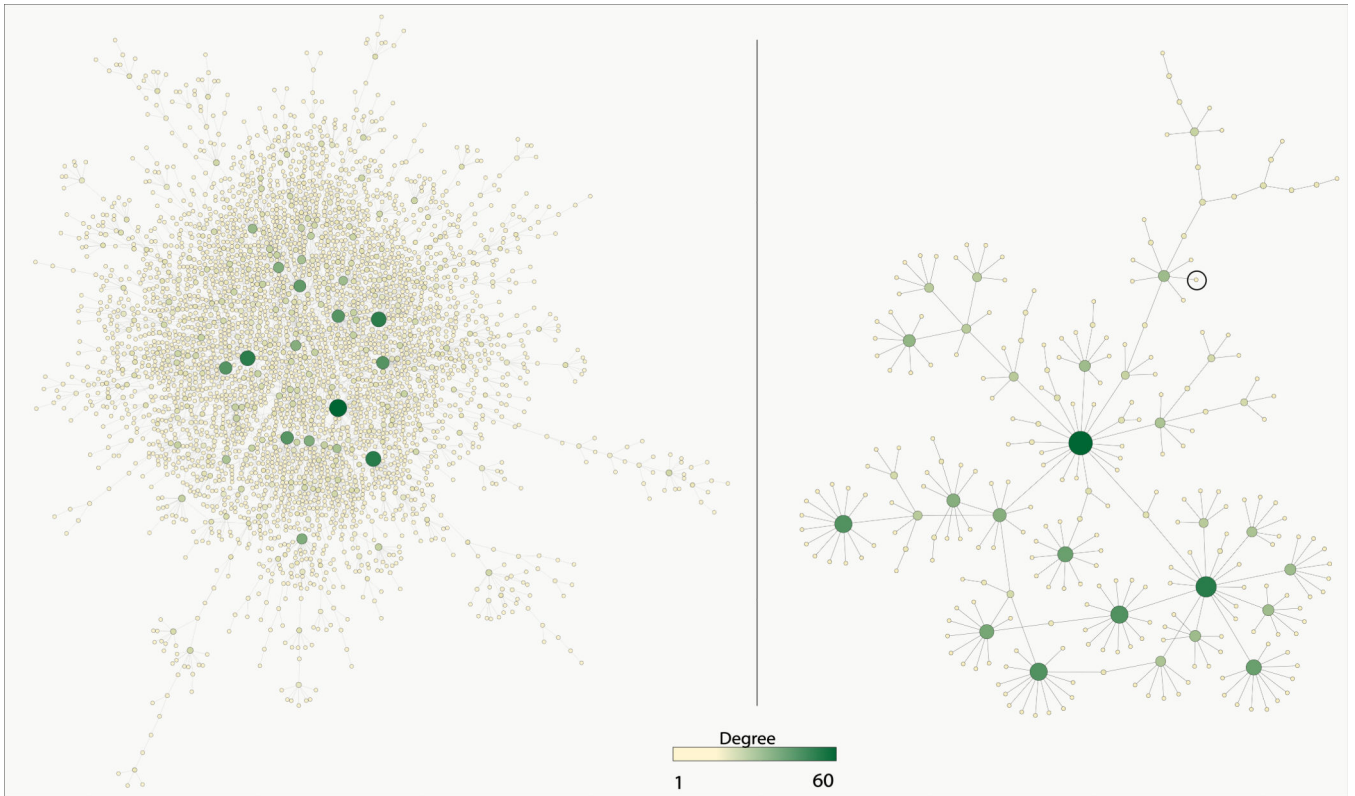


Fig. 6. Single run of largest component (left) and one eight-step walk from a randomly chosen node (right)

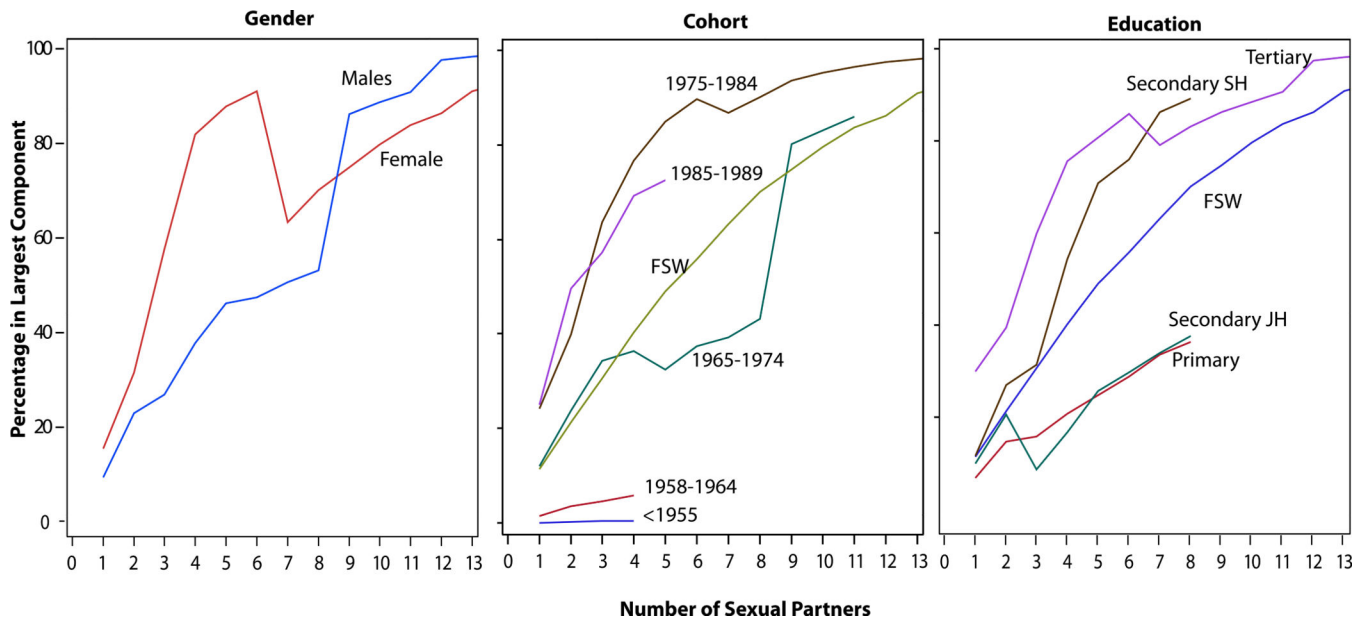


Fig. 7.
Node membership in largest component by demographic attributes and degree

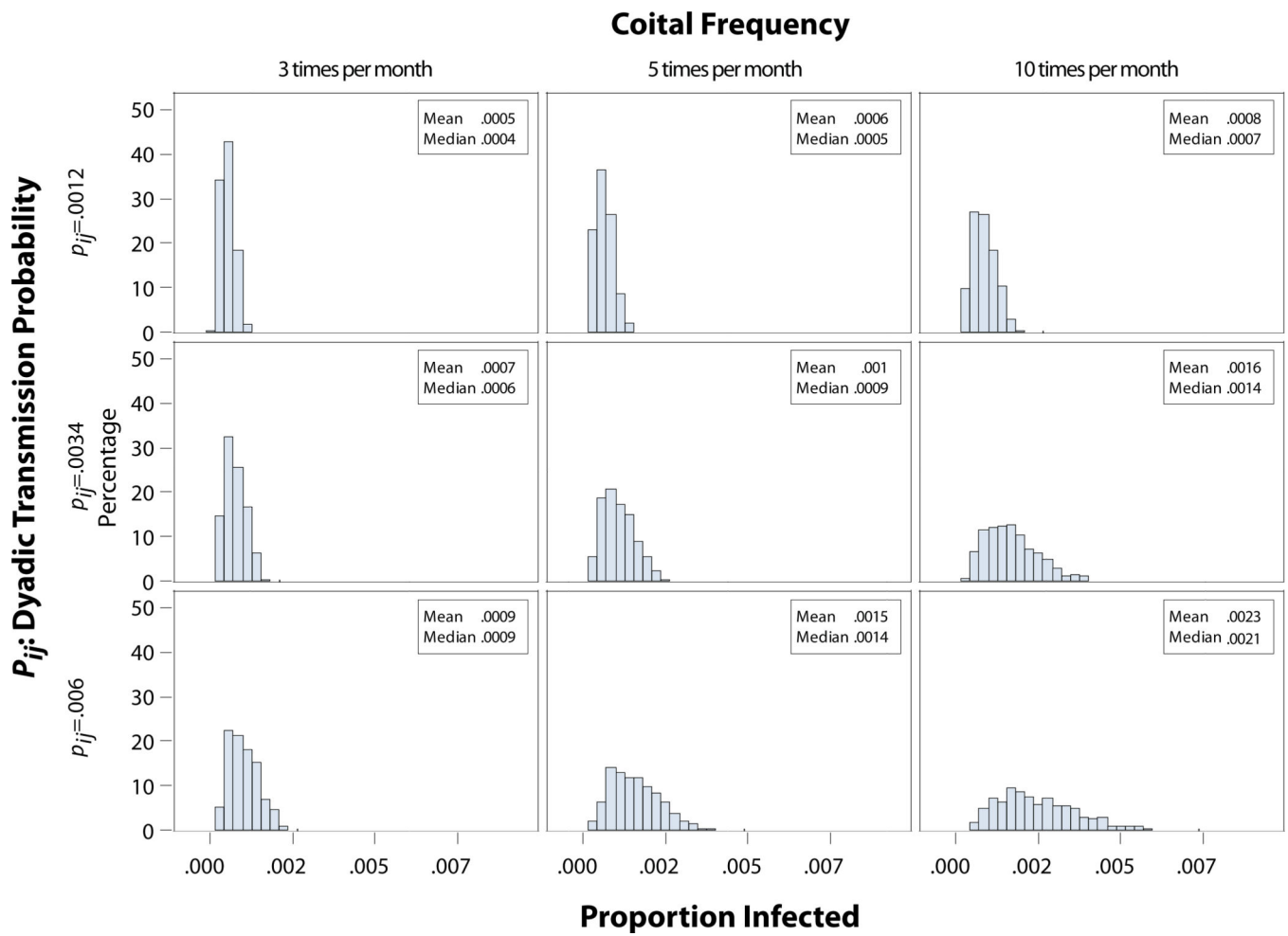


Fig. 8. Proportion of the population infected, by dyadic transmission probability and coital frequency, based on dynamic network disease transmission simulation ($N = 1,200$ in each cell)

Table 1

Biological and behavioral parameter values for the model of HIV transmission

Parameter				
Biological, All Types of Partnership				
Infectivity of HIV per sexual act (β_a)	0.0012 ^a	0.0034 ^b		0.006 ^c
Behavioral by Type of Partnership				
	Non-FSW ties		FSW ties	
Condom use (% always condom protected) (c)	12.5 ^d	30	38	22 ^d 50 53
Coital frequency per month (n)	3	5 ^e	10	1 3 5

^a Average of M-F and F-M HIV transmission probabilities per vaginal coital act in high-income countries. These are frequently cited probabilities of transmission of HIV per single unprotected act estimated from a strictly controlled study of 525 HIV-discordant European couples (Downs and De Vincenzi 1996).

^b Average of M-F and F-M HIV transmission probability per sexual act in low-income country <http://www.aidsmap.com/Estimated-risk-per-exposure/page/1324038/>

^c Estimate of late-stage HIV infectivity (vaginal sex). *Source:* <http://www.aidsmap.com/Estimated-risk-per-exposure/page/1324038/>

^d *Source:* SSNS. Non-FSW ties condom use: average male and female respondents reporting always using condoms with marital partners. FSW ties condom use: male respondents reporting always using condoms with FSW's partners.

^e *Source:* SSNS.