Journal of the Royal Statistical Society
SERIES A Statistics in Society

J. R. Statist. Soc. A (2012)
175, *Part* 1, *pp.* 191–216

The sensitivity of respondent-driven sampling

Xin Lu,

Karolinska Institutet, Stockholm, and Stockholm University, Sweden

Linus Bengtsson,

Karolinska Institutet, Sweden

Tom Britton,

Stockholm University, Sweden

Martin Camitz,

Karolinska Institutet, Stockholm, Sweden

Beom Jun Kim,

Sungkyunkwan University, Suwon, Republic of Korea

Anna Thorson

Karolinska Institutet, Stockholm, Sweden

and Fredrik Liljeros

Stockholm University, Sweden

[Received February 2010. Final revision March 2011]

Summary. Researchers in many scientific fields make inferences from individuals to larger groups. For many groups, however, there is no list of members from which to draw a random sample. Respondent-driven sampling (RDS) is a relatively new sampling methodology that circumvents this difficulty by using the social networks of the groups under study. The RDS method has been shown to provide unbiased estimates of population proportions given certain conditions. The method is now widely used in human immunodeficiency virus related studies among high risk populations globally. We test the RDS methodology by simulating RDS studies on the social networks of a large Lesbian, gay, bisexual and transgender Web community. The robustness of the RDS method is tested by violating, one by one, the conditions under which the method provides unbiased estimates. Simulations indicate that the bias is large if networks are directed or respondents choose to invite people on the basis of characteristics that are correlated with the study outcomes. The bias and variance increase if participants invite close as opposed to more distant friends whereas sampling in denser networks sharply reduces variance. However, the RDS method shows strong resistance to sampling without replacement, low response rates and certain errors in the participants' reporting of their network sizes, as well as the selection criteria of seeds. The effects of network structure and the number of seeds and coupons are also discussed.

Keywords: Directed network; Hidden population; Network; Respondent-driven sampling; Sampling; Sensitivity

Address for correspondence: Xin Lu, Department of Sociology, Stockholm University, SE-106 91, Stockholm, Sweden.
E-mail: lu.xin@sociology.su.se

 0964–1998/12/175191

1. Introduction

Hidden or difficult-to-reach populations, such as injecting drug users, men who have sex with men and sex workers, are generally difficult to access because of their strong privacy concerns and a lack of a well-defined sampling frame from which a random sample can be drawn (Heckathorn, 1997). Sampling frames are also lacking for many groups without strong privacy concerns, such as jazz musicians (Heckathorn and Jeffri, 2001). Methods for obtaining information about such groups have involved contacting especially knowledgeable people within the group, which is known as key informant sampling (Deaux and Callaghan, 1985), targeted sampling where participants are recruited from locations where group members are known to pass (Watters and Biernacki, 1989) or snowball sampling where members of a group are asked to give the researchers contact details of others in the same group (Erickson, 1979). However, these methods all introduce a considerable selection bias, which impairs generalization of the findings from the sample to the population studied (Heckathorn, 1997; Magnani *et al.*, 2005).

Respondent-driven sampling (RDS) is a method which was developed to overcome the challenges of selection bias when sampling hidden populations (Heckathorn, 1997, 2002; Salganik and Heckathorn, 2004; Salganik, 2006; Volz and Heckathorn, 2008). An RDS study starts out by purposively selecting some participants who are members of the study population (usually 5–15). These people are called 'seeds'. The seeds are given a number of invitation coupons (usually 3) to distribute to friends and acquaintances within the study population. If those friends who receive a coupon decide to participate, they are in turn given the same number of coupons to invite further participants. Participants are rewarded for their personal participation in the study, as well as for each peer they invite and who also participates. The invitation coupon contains a serial number that enables the researchers to follow the recruitment chains in the sample. If the recruitment chains are sufficiently long, the sample composition stabilizes and becomes independent of the characteristics of the seeds. Additionally, each participant is asked for the number of people he or she knows within the study population, known as his or her 'personal network size' or 'degree'. The degree of a participant is important to collect as participants with large degrees are oversampled and participants with small degrees are undersampled. Knowing the degree of each participant hence allows adjustment for this bias.

When the sample has been collected, the proportion of people with the characteristic A in the population can be estimated by the updated RDS estimator *RDSII* (Volz and Heckathorn, 2008):

$$\hat{P}_A = \sum_{i \in A \cap S} d_i^{-1} \Big/ \sum_{i \in S} d_i^{-1} \tag{1}$$

where d_i is the degree of individual i, and S the set of sampled individuals.

Volz and Heckathorn (2008) proved that RDSII provides asymptotically unbiased estimates if the following assumptions are fulfilled.

(a) Reciprocity: individuals in the studied population maintain and recruit peers through reciprocal relationships, i.e. the network within which recruitment happens is undirected.
(b) Connectedness: each individual in the population studied has a chance of being invited to participate, i.e. the network forms a single component.
(c) Sampling is with replacement: individuals are allowed to be recruited into the sample more than once.
(d) Degree: respondents can accurately report their degree in the network.
(e) Random recruitment: peer recruitment is a random selection from the respondents' personal network.
(f) Each respondent recruits a single peer, i.e. the number of recruitment coupons is 1.

The ability to produce population estimates and a feasible field implementation have contributed to a rapid increase in RDS studies conducted globally in recent years. To date, well over 100 studies in over 30 countries have been performed (Johnston *et al.*, 2008; Malekinejad *et al.*, 2008).

However, the assumptions underlying the RDS estimator are not easily met in real life. First, most social networks contain directed edges, or edges that do not have the same strength in both directions. Second, to prevent participants from colluding to recruit each other back and forth to gain rewards, real life RDS studies sample without replacement, meaning that respondents can participate only once. Third, it is difficult for respondents to report their degree accurately (Marsden, 2005). Fourth, participants usually pass their coupons to peers with whom they have a close rather than a more distant relationship, which is not a random selection (Wang *et al.*, 2005; Frost *et al.*, 2006). Fifth, to avoid recruitment chains stopping too early, researchers most often use three coupons rather than one (Johnston *et al.*, 2008; Malekinejad *et al.*, 2008).

How well the theoretical assumptions are fulfilled in real life studies and whether deviations from these assumptions critically affect RDS estimates have been discussed in the literature (Heckathorn, 1997, 2002; Salganik and Heckathorn, 2004; Heimer, 2005; Salganik, 2006; Volz and Heckathorn, 2008; Goel and Salganik, 2009, 2010; Gile and Handcock, 2010). Goel and Salganik (2010) tested the performance of RDS on a high risk heterosexual network and on school friendship networks. They found that the design effect (the variance of the RDS estimates divided by the variance under simple random sampling) was much larger than previously assumed. The most recent and comprehensive study of violations of the theoretical assumptions (Gile and Handcock, 2010) simulated RDS studies on artificial networks constructed from pilot study data from the US Centers for Disease Control and Prevention surveillance programme (Abdul-Quader *et al.*, 2006). Gile and Handcock (2010) analysed bias induced by violations of assumptions (c), (e) and (f) in the context of a total population of 1000 people and a sample size ranging from 500 to 950 participants (sampling fraction: 50–95%). They addressed the possibility of a reduction of bias by discarding early waves and found a potential bias caused by preferential selection of peers and sampling without replacement. However, the numbers of seeds, coupons and waves were fixed and many other assumptions that might affect the RDS estimates, such as directness of networks, recruitment failures and degree reporting error, were not simulated.

On the basis of current literature and the increasing use of RDS within research, we thus identified a need for systematically testing the robustness of the RDS method when sampling diverges from the basic assumptions in the analytical proof. This study simulates RDS studies on a real on-line network of 16082 homosexual men, as well as on several variants of this network. In these simulations we systematically and to varying extent violate each assumption one by one and describe the effects on the RDS estimates.

We use RDSII for all RDS estimates in this paper as this estimator has improved analytical power compared with earlier RDS estimators and provides equivalent estimates when data smoothing is used (Heckathorn, 2007; Volz and Heckathorn, 2008; Gile and Handcock, 2010)

Four measurements are used throughout this paper: the average estimate (AE), which is defined by the mean of the RDSII estimates, $\mathrm{AE}_j = \sum_{i=1}^{m} \mathrm{est}_{ij}/m$, where est_{ij} is the estimate of RDSII at the ith simulation when the sample size is j; the bias, which is defined by the absolute difference between AE and the true population, $\mathrm{Bias}_j = |\mathrm{AE}_j - P^*|$; the standard deviation (SD) of estimates for a given sample size j, SD_j; and finally the mean absolute error (MAE) of estimates for a given sample size j, $\mathrm{MAE}_j = \sum_{i=1}^{m} |\mathrm{est}_{ij} - P^*|/m$. An alternative measure would be to use the root-mean-square error. Here we have used the MAE to reduce the influence of outliers.

The rest of this paper is organized as follows: in Section 2, we give a brief description of our data and networks; in Section 3, we describe the results of simulating RDS studies in the networks when all the assumptions are satisfied; in Section 4, we test the effects of violating the assumptions one by one; in Section 5, we summarize and draw our conclusions.

2. The men-with-men network

2.1. Data collection

'Qruiser' (`http://www.qx.se`), is the Nordic region's largest and most active Web community for homosexual, bisexual, transgender and 'queer' people. Contacts between members on the Web site are maintained mainly by a 'favourites list', on which each member can add any other member without approval from that member. Members can attend clubs (Web pages with specific topics) and send messages to each other (Rybski *et al.*, 2009).

We collected information on personal profiles as well as on all messages that were sent within the Web community from December 15th, 2005, to January 18th, 2006. During the 63 days of this data collection period, 12590911 messages were recorded and 184819 distinct members were registered on the Web site.

2.2. Network formation

On the basis of the membership profiles, we extracted a network that contained only members characterizing themselves as homosexual males. We define an outgoing edge to be formed if a member has another member on his favourites list. An edge is called reciprocal if a connected pair of members have both an ingoing and an outgoing edge between each other. If a pair does not have both an ingoing and an outgoing edge between each other, it is called irreciprocal. To avoid the inclusion of inactive people, members were required to have sent at least one message during the data collection period.

For our research purpose, only members of the giant connected component (GCC), which is defined by the largest component connected with only reciprocal edges, were kept as nodes (16082 active, gay men). By keeping only the reciprocal edges in the GCC, we obtained an undirected network (G1), with an average degree of 6.74. By keeping both reciprocal and irreciprocal edges in the GCC, we obtained a directed network (G2) with an average degree of 17.2. Note that the definition of the GCC ensures that all nodes have a chance of being recruited with RDS sampling in both G1 and G2. Degree distributions for both G1 and G2 are plotted in Fig. 1. The distributions are very skewed. For instance, half of the members in G2 have no more than 10 outgoing edges, whereas a small proportion of members have a large number of outgoing edges.

2.3. Homophily

An important issue for chain referral sampling is the homophily of edge formation. Homophily is the probability that participants connect with friends who are similar to themselves rather than connecting randomly (Rapoport, 1980; Morris and Kretzschmar, 1995; McPherson *et al.*, 2001; Heckathorn, 2002). The homophily, which is defined in accordance with Heckathorn (2002), of different groups in our network is shown in Table 1. The homophily with respect to *age* and *county* is fairly large, indicating that a fair part of the edges were formed between members of the same age or between members living in the same county. Taking county within the undirected network G1 as an example, members who live in Stockholm formed edges with members who also lived in Stockholm 50% of the time, whereas they formed edges randomly among all

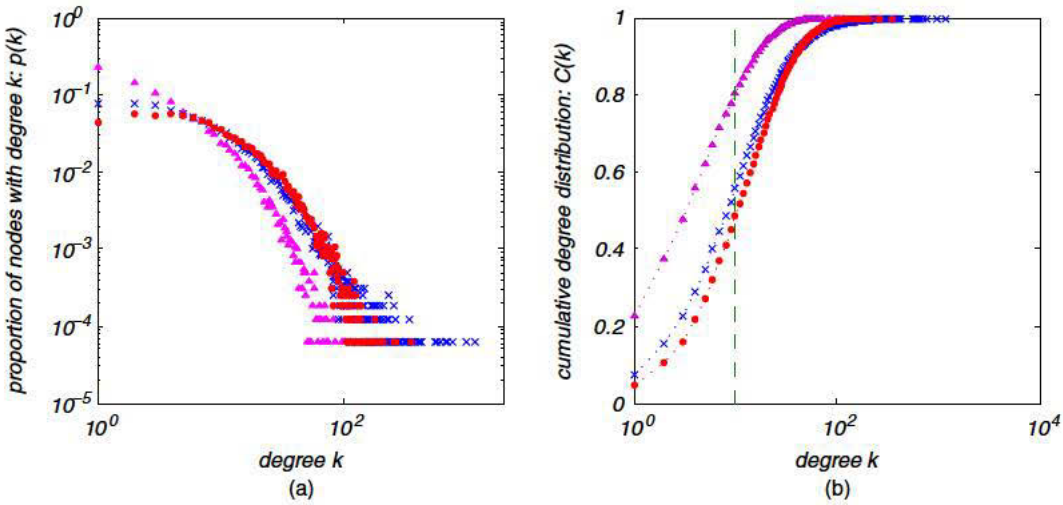


Fig. 1. (a) Degree distribution and (b) cumulative degree distribution (the out-degree is the number of outgoing edges that leave a node and the in-degree is the number of ingoing edges that point to the node): ▲, degree (G1); ×, out-degree (G2); ●, in-degree (G2)

Table 1. Proportions P^*, homophilies H and activity ratio w of groups in the undirected (G1) and directed (G2) networks

	Results for the following groups:							
	Age		*County*		*Civil status*		*Profession*	
	Before 1980	*Others*	*Stockholm*	*Others*	*Single*	*Others*	*Employed*	*Others*
P^*(%)	77.77	22.23	38.79	61.21	40.39	59.61	38.19	61.81
H G1	0.40	0.37	0.50	0.40	0.05	0.08	0.13	−0.05
H G2	0.23	0.34	0.50	0.28	0.03	0.07	0.06	0.02
w G1	1.05	0.95	1.22	0.82	0.97	1.03	1.21	0.83
w G2	1.22 (0.95)	0.82 (1.05)	1.15 (1.32)	0.87 (0.76)	0.98 (0.96)	1.02 (1.04)	1.10 (1.05)	0.91 (0.95)

cities (including Stockholm) the remaining 50% of the time. The homophily for people living in Stockholm is thus 0.5. Homophilies for *civil status* and *profession* are very small, indicating that edges were formed as if members, regarding civil status and profession, chose randomly among other members.

2.4. Activity ratio

The activity ratio is the mean degree of the group of interests divided by the mean degree of the group of non-interest: $w = \bar{d}_A / \bar{d}_{S-A}$. Gile and Handcock (2010) have shown that, when the sample constitutes a large fraction of the population studied, subtle changes in w might result in large bias. The activity ratios for our networks are listed in Table 1. The activity ratios for the directed network are calculated on the basis of the out-degree (the in-degree is shown in parentheses). As we can see, on the undirected network, i.e. G1, the activity ratios are rather low for age and civil status, whereas country and profession have relatively large values, indicating

that individuals living in Stockholm, or individuals who are employed, have on average 1.2 times larger social networks than people living outside Stockholm or people who are unemployed. The former are thus more likely to be recruited into the sample.

2.5. Network variation

To avoid misleading conclusions resulting from the effects of network structure and edge density in our simulations on the undirected network (G1), we created two variants of G1: the first type of networks ($G1_{add}$) was obtained by randomly adding reciprocal edges with properties proportional to G1 until the average degree was increased by 20, for each property, separately. The second type of networks ($G1_{rand}$) was obtained by randomly rewiring each pair of reciprocal edges to another node with the same property as the former node. After the procedures above, we obtained four dense networks and four randomized networks, all with the homophily unchanged for each property respectively. The details of the procedures for generating these networks can be found in Appendix A. Their degree distributions are shown in Fig. 2.

To test the effects of preferential recruitment in RDS (Section 4.4), we weighted each reciprocal edge in G1 in two ways: by the maximum number of sent messages in any one direction, and by the minimum number of sent messages in any one direction. For example, if node i sent 10 messages to node j and received five back from j, the weight on edge $e_{i,j}$ (or $e_{j,i}$) would

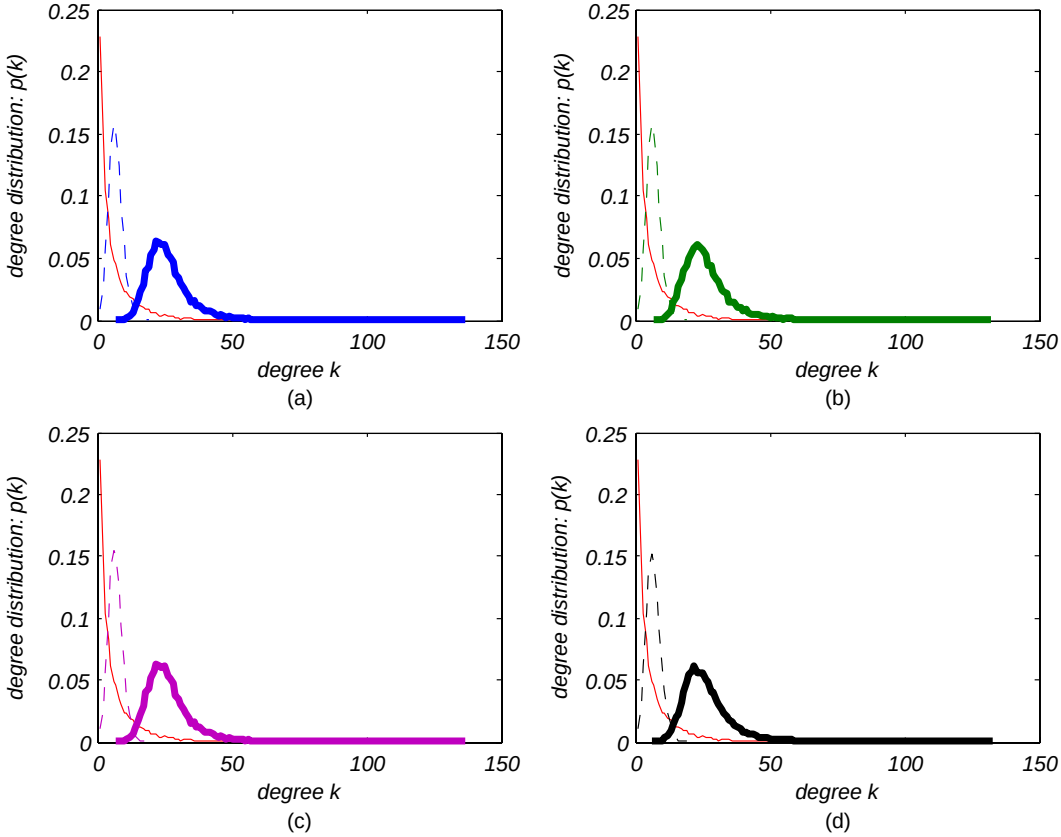


Fig. 2. Degree distributions for G1 (———), $G1_{add}$ (━━━) and $G1_{rand}$ (– – –): (a) age; (b) county; (c) civil status; (d) profession

be 10 for the maximum weighted network ($G1_{max}$) and five for the minimum-weighted network ($G1_{min}$). In these two weighted networks, respondents were supposed to recruit peers with probability proportional to the edge weights.

We now proceed to describe the results of simulated RDS samplings under varying circumstances in the networks that were described above. We compare the true population proportions of the four variables in Table 1 (two with high homophily and two with low) with the RDS estimates given by the simulated samplings. Unless otherwise stated, respondents were set to use all their coupons (or to recruit all their friends if their out-degrees were too low), and all simulations were repeated 10000 times.

3. Respondent-driven sampling on the undirected network

We first ran simulations on the undirected network G1 to see whether RDSII worked well when all the stated assumptions (a)–(f) were satisfied. We started each simulation with a single randomly selected seed and we restricted the number of coupons to 1, i.e. each recruiter could recruit only one other person (assumption (f)). All respondents were selected randomly from the recruiters' personal networks (assumption (e)), and nodes could be selected multiple times

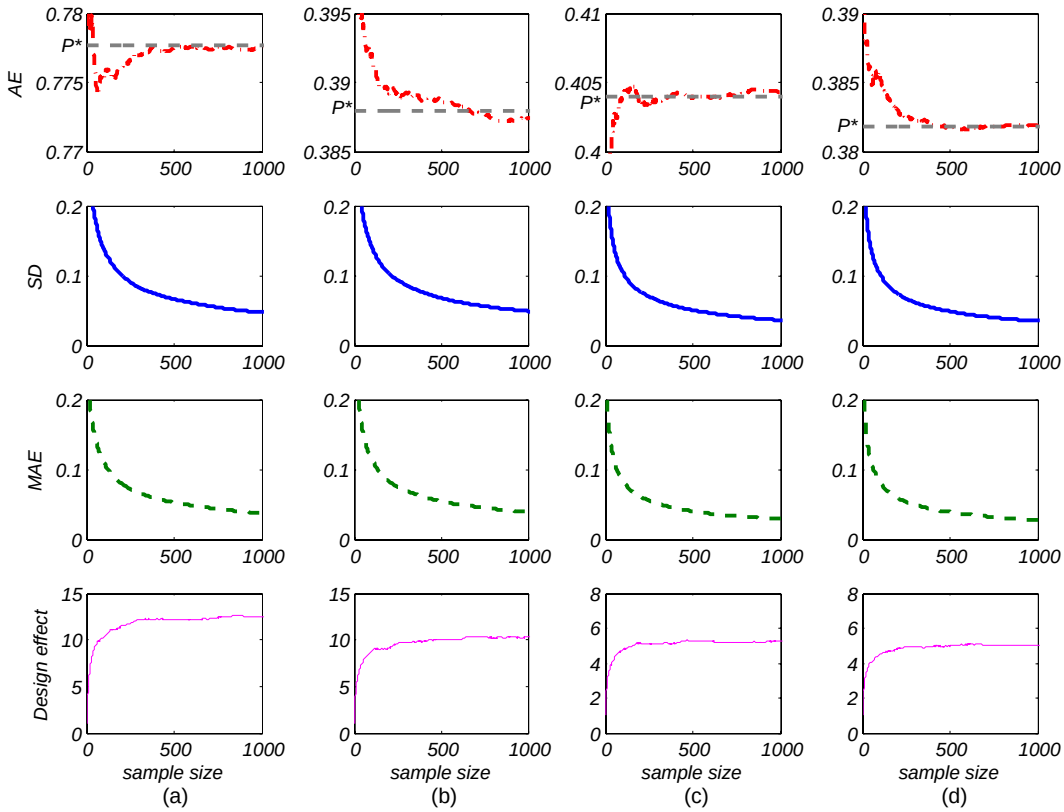


Fig. 3. RDSII estimations on the undirected network G1 (the average estimates approached the true proportions very fast; when the sample size was 500, the bias was only 0.0002, 0.0009, 0.00002 and 0.0002 for age, county, civil status and profession respectively; design effects are substantial for age and county (homophily 0.4/0.37 and 0.5/0.4) and lower for civil status and profession (homophily 0.05/0.08 and 0.13/−0.05)): (a) age; (b) county; (c) civil status; (d) profession

(sampling with replacement; assumption (c)). Since all participants' degrees were assumed to be known by the participants themselves, and G1 constitutes a single connected component with only reciprocal edges, assumptions (a), (b) and (d) were also satisfied. We kept recruiting participants in the simulation until the sample size reached 10000 participants. The AE, SD, MAE and design effects are shown in Fig. 3 for sample sizes of less than 1000 (see Fig. 4 for sample sizes above 1000). Even though our network is sparse compared with reported studies (Heckathorn and Jeffri, 2003; Ramirez-Valles *et al.*, 2005; de Mello *et al.*, 2008; Volz and Heckathorn, 2008), the RDSII estimates converged to the true population proportions P^* very quickly.

The SD was around 0.05, and the MAE was around 0.04 when sample sizes were between 500 and 1000 participants. The SD and MAE decreased to 0.02 when the sample size approached 10000 (see Fig. 4). The design effects of RDSII under fully satisfied assumptions were around 13 and 10, for age and county respectively, and 5 for both civil status and profession. These values are much larger than what has been assumed earlier but are consistent with what Goel and Salganik (2010) found in their networks. It is worth noting that when all the assumptions are fulfilled the design effect above a sample size of 100 is almost constant, telling us that with increasing sample size the variance of RDSII decreases at the same speed as under simple random sampling.

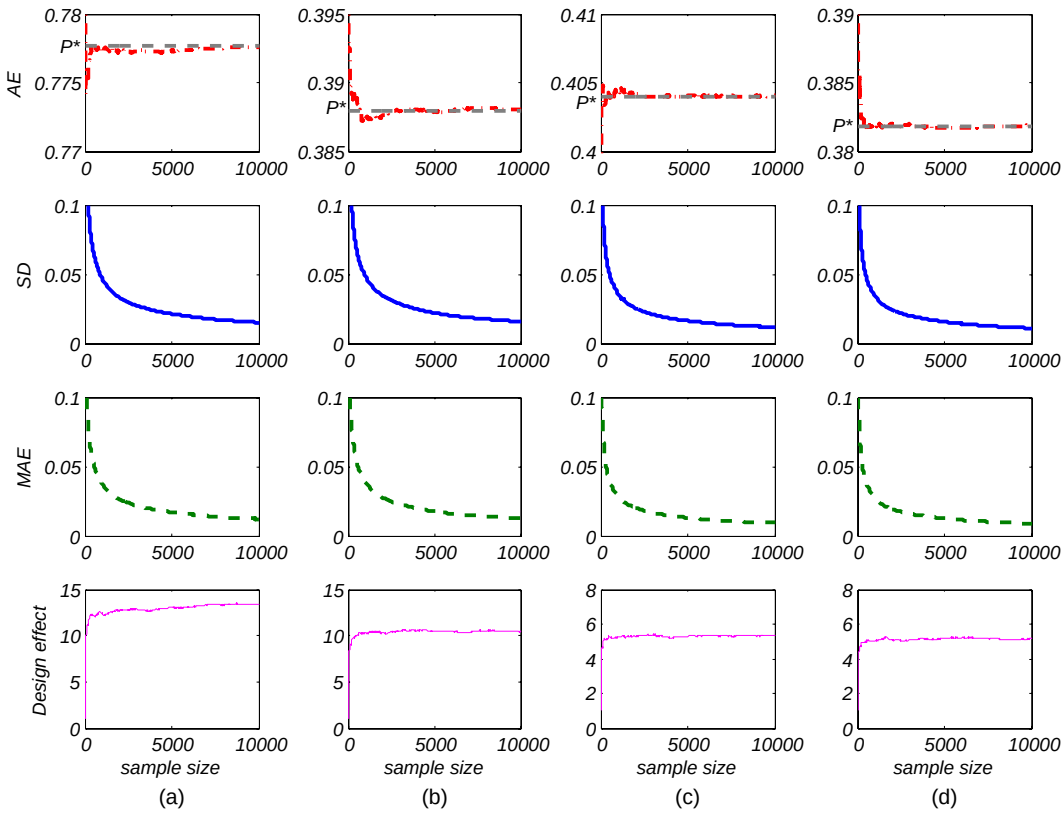


Fig. 4. RDSII estimations on the undirected network G1, for sample size larger than 1000: (a) age; (b) county; (c) civil status; (d) profession

In summary, the average RDSII estimates (the AEs) on the undirected network with all assumptions fulfilled were unbiased and converged rapidly to the true population proportion. The design effects were, however, substantial.

4. Violations of assumptions

4.1. Respondent-driven sampling on directed networks

If a directed network forms a giant strongly connected component (Schwarte *et al.*, 2002) in which every node can be reached by any other, and assumptions (c) and (d) are satisfied, the next selected individual in an RDS depends only on the current respondent, which is known as the Markov property (Hastings, 1970). Hence the RDS can be modelled as a Markov process with the following transition matrix (Heckathorn, 2007; Volz and Heckathorn 2008):

$$A = \begin{pmatrix} 0 & e_{12}/d_1^{\mathrm{o}} & \dots & e_{1N}/d_1^{\mathrm{o}} \\ e_{21}/d_2^{\mathrm{o}} & 0 & \dots & e_{2N}/d_2^{\mathrm{o}} \\ \vdots & \vdots & \ddots & \vdots \\ e_{N1}/d_N^{\mathrm{o}} & e_{N2}/d_N^{\mathrm{o}} & \dots & 0 \end{pmatrix} \tag{2}$$

where $e_{ij} = 1$ if there is an edge from individual i to individual j, and $e_{ij} = 0$ otherwise, and d_i^{o} is the out-degree of i. The equilibrium state distribution for this process is a vector $X = (x_1, x_2, \dots, x_n)^{\mathrm{T}}$ such that

$$X^{\mathrm{T}} A = X^{\mathrm{T}}. \tag{3}$$

If the out-degree and in-degree are equal for all nodes it can be verified that

$$x_i = d_i \Big/ \sum_{j=1}^{N} d_j. \tag{4}$$

Actually, equation (4) is the underlying equation from which RDSII is derived (Volz and Heckathorn, 2008; Goel and Salganik, 2009).

We now show how the equilibrium solution may be used in estimations from RDS samples in directed networks. We can rewrite equation (3) as

$$A^{\mathrm{T}} X = X \tag{5}$$

and note that X should be an eigenvector with eigenvalue 1 for A^{T} since the eigenvalues λ are defined by $A^{\mathrm{T}} X = \lambda X$. The existence of an eigenvalue 1 for A^{T} can be easily proved as the all 1 vector is an eigenvector for A (Woess, 1994; Page *et al.*, 1999).

We let $V = (v_1, v_2, \dots, v_N)$ be the normalized eigenvector of A^{T} for eigenvalue 1. The members of an RDS sample $\{s_1, s_2, \dots, s_n\}$ can then be weighted by the reciprocals of their values in V so that an estimate of the proportion of individuals in group A in the population is

$$\hat{P}_A = \sum_{s_i \in A} \frac{1}{v_{s_i}} \Big/ \sum_{j=1}^{n} \frac{1}{v_{s_j}}. \tag{6}$$

We denote the estimator of equation (6) from RDS samples in a directed network using these weights as 'eig' to weight the RDS samples in a directed network. Note that we can barely know the network value V from an RDS sample.

Both RDSII and eig estimations on the directed network (G2) are presented in Fig. 5. Not surprisingly, the RDSII estimates were biased for all groups. For age and county, these biases

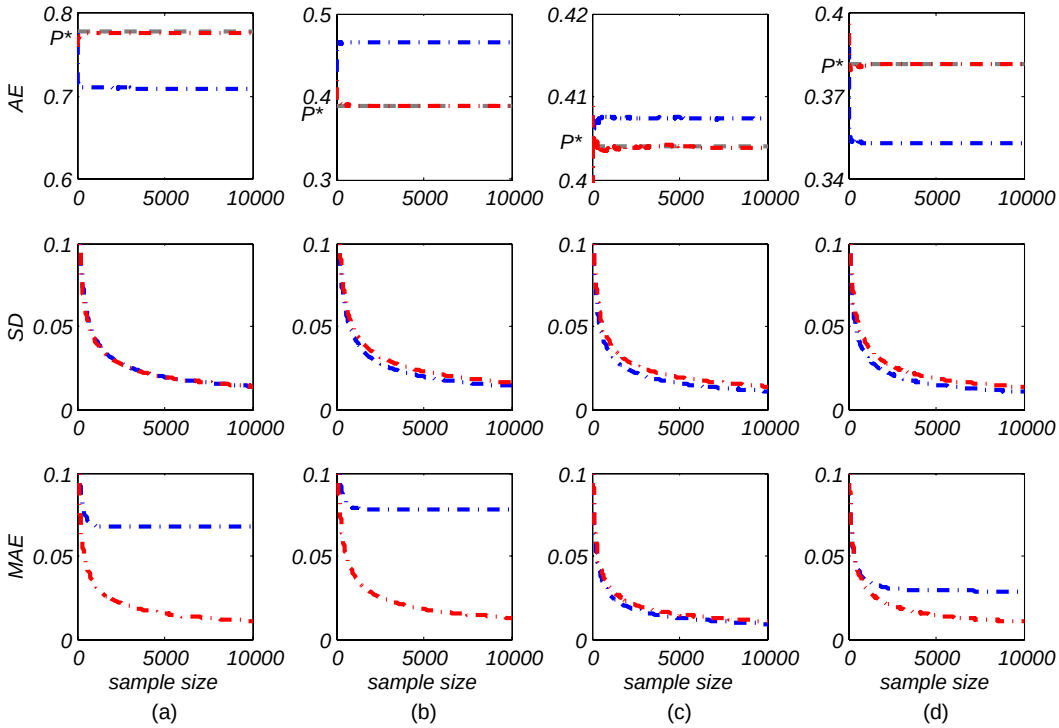


Fig. 5. Estimations on the directed network G2 (number of seeds, 1; numbr of coupons, 1; with replacement; -·-·-, RDSII estimates with out-degree; -··-··-, estimates weighted by eigenvectors): (a) age; (b) county; (c) civil status; (d) profession

were as high as 0.06, whereas civil status and profession performed better, at 0.005 and 0.022 respectively. However, the eig estimates weighted by equation (6) agreed well with the true population proportions.

The SDs were similar for all four groups (and very similar to the SD of the undirected networks). However, the MAE of RDSII in the directed network was much higher than that of the undirected networks for age and county (0.07–0.08), indicating that, if the network under study is partly directed, the use of RDSII estimations could result in relatively large errors. We can see that the MAE for civil status and profession were small, telling us that directness of edges might have little effect on RDSII for groups with low homophily.

4.2. Sampling without replacement

It is generally believed that sampling without replacement creates negligible bias compared with sampling with replacement in RDS when the sample size is small relative to the population (Heckathorn 1997, 2002; Volz and Heckathorn, 2008). We tested this proposition in our undirected network. To increase the generalizability of our results, we also compared the replacement effect on $G1_{add}$ and $G1_{rand}$. Results are shown in Fig. 6 for sample sizes smaller than 1000 and in Fig. 7 for sample sizes up to 10000. For large sample sizes the RDSII estimations for sampling with replacement were almost unbiased on all networks. This result held up even when the sample occupied a large fraction of the whole population. The estimations for sampling without replacement in contrast were biased in different directions in the different networks. For sample sizes below 1000 there was no clear trend and the average estimates of sampling

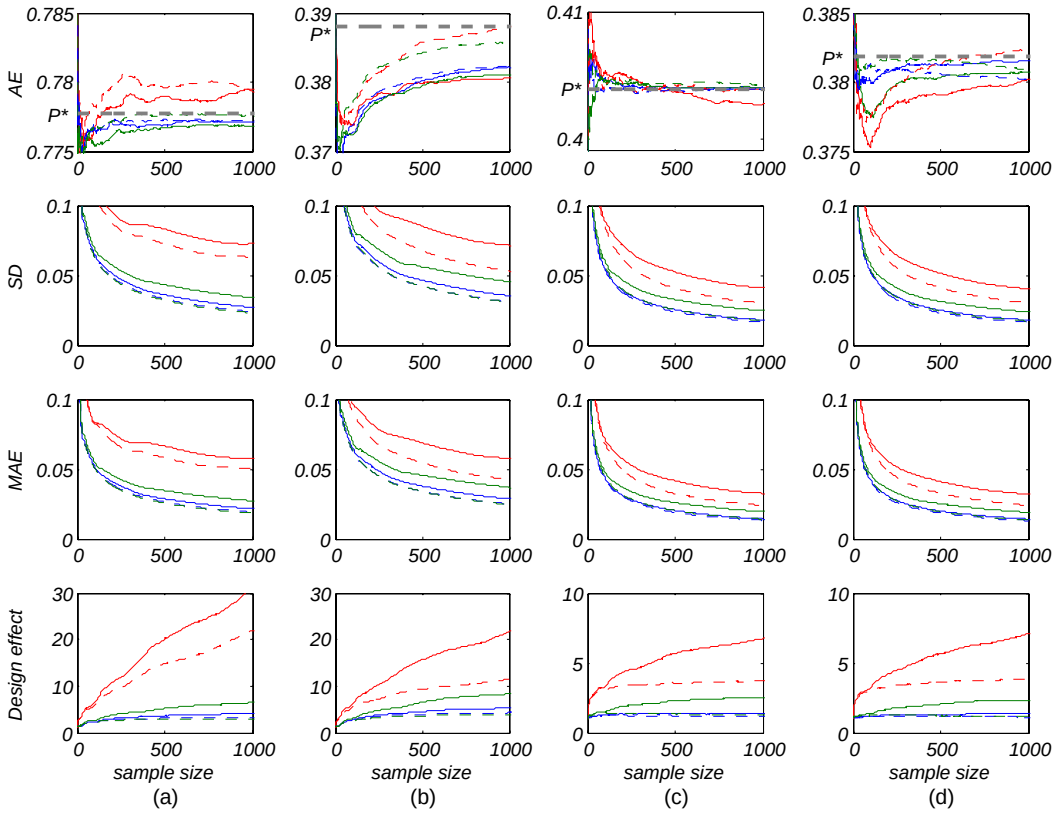


Fig. 6. Effects of network structures and replacement (number of seeds, 1; number of coupons, 3; ———, G1, sampling with replacement; – – –, G1, sampling without replacement; ———, $G1_{add}$, sampling with replacement; – – –, $G1_{add}$, sampling without replacement; ———, $G1_{rand}$, sampling with replacement; – – –, $G1_{rand}$, sampling without replacement; seeds were randomly selected at the beginning of each simulation): (a) age; (b) county; (c) civil status; (d) profession

without replacement were sometimes closer to the true population than those for sampling with replacment (see Fig. 6).

The sampling without replacement estimates always have smaller SD and MAE than those for sampling with replacement. This is especially apparent in G1. Simulations that are not included in this paper indicate that networks with skewed degree distributions result in larger variances than those with a Poisson distribution and, as networks become denser, the variance becomes smaller.

We can see that when the RDS study starts with 10 seeds and three coupons, given that all the other assumptions are fulfilled, the design effects for all the four properties increased with the sample size. This is most evident for sampling with replacement in G1 and $G1_{rand}$. For sample sizes of 500, the design effect for age increased to 22 and for country to 15 in G1. However, consistent with what we observed from the SD of estimates from sampling without replacement in Fig. 6, the design effects of sampling without replacement are all smaller than for sampling with replacement for all variables. For civil status and profession, they are even smaller than the ideal situation for RDSII when all assumptions are fulfilled. A likely mechanism through which sampling without replacement decreases the design effect is forcing nodes to recruit new nodes rather than already sampled nodes, thereby enabling the recruitment chains

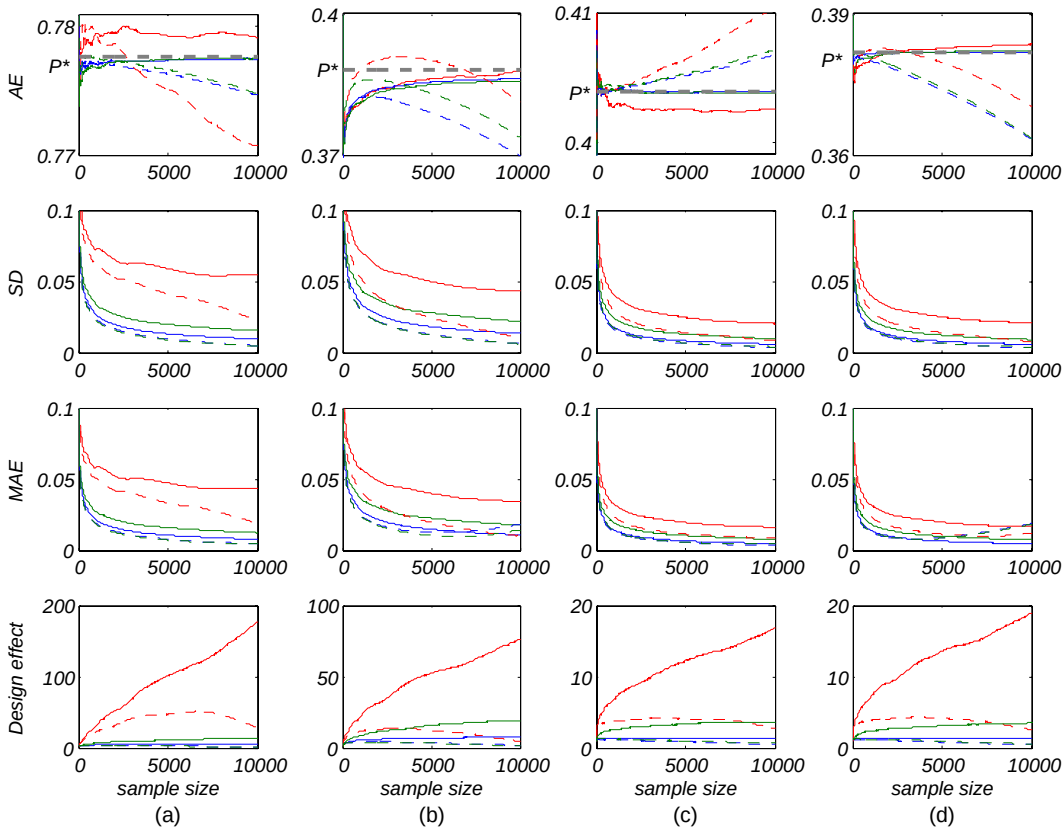


Fig. 7. Effects of network structures and replacement, for sample size larger than 1000 (number of seeds, 1; number of coupons, 3; ———, G1, sampling with replacement; – – –, G1, sampling without replacement; ———, $G1_{add}$, sampling with replacement; – – –, $G1_{add}$, sampling without replacement; ———, $G1_{rand}$, sampling with replacement; – – –, $G1_{rand}$, sampling without replacement; seeds were randomly selected at the beginning of each simulation): (a) age; (b) county; (c) civil status; (d) profession

to penetrate new areas of the network. This will result in a more diverse set of nodes, which are more likely to represent the entire network than if recruitment chains repeatedly sample already explored areas. The lower variance found under sampling without replacement in this study is also consistent with the findings of Gile and Handcock (2010). Differences between sampling without replacement and sampling with replacement are smallest for $G1_{add}$, supporting the intuition that dense networks are less affected by replacement or non-replacement sampling.

4.3. Rejecting invitations and forgetting peers

Although a great majority of participants in RDS studies report a social network size larger than 3, not all distributed coupons result in study participations (Johnston *et al.*, 2008; Malekinejad *et al.*, 2008; de Mello *et al.*, 2008). This could be seen as a violation of assumption (d), which states that participants can accurately report their personal network size, which is assumed to reflect the chances of being invited or of the number or peers who have a chance of being invited by the person, and assumption (f), which states that all participants use their one coupon to make one successful recruitment. The latter assumption includes the dual assumption that each

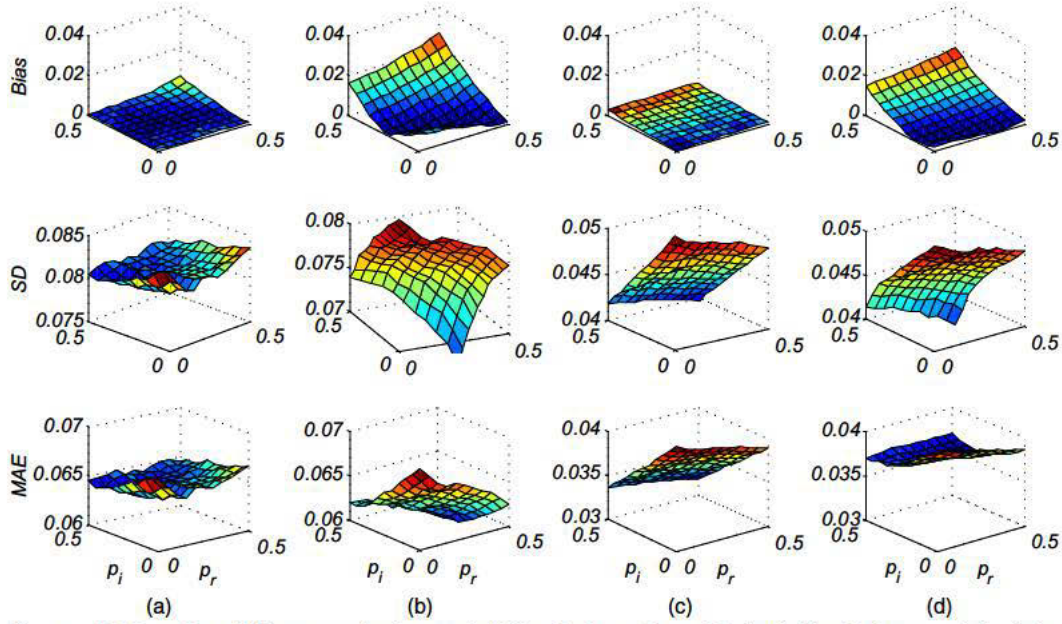


Fig. 8. RDS on G1 with ignore and reject probabilities independent of the individuals' characteristics (simulations were repeated 100000 times for each combination; number of seeds, 10; number of coupons, 3; sampling with replacement; sample size 500; seeds were randomly selected at the beginning of each simulation): (a) age; (b) county; (c) civil status; (d) profession

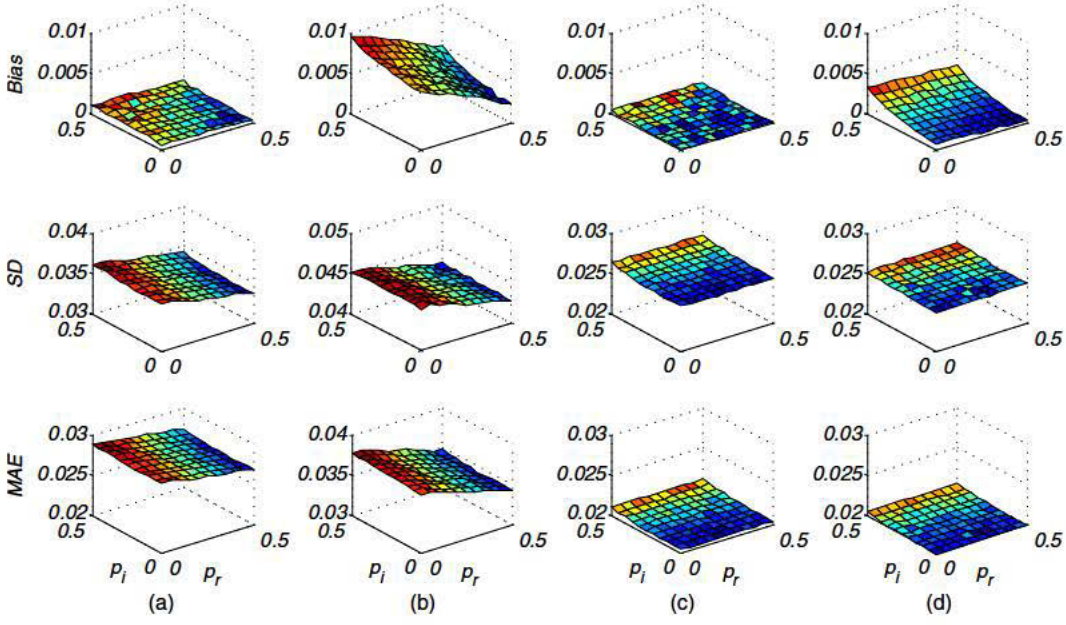


Fig. 9. RDS on $G1_{add}$ with ignore and reject probabilities independent of the individuals' characteristics (simulations were repeated 100000 times for each combination of probabilities number of seeds, 10; number of coupons, 3; sampling with replacement; seeds were randomly selected from the beginning of each simulation; estimates were calculated when sample sizes reached 500): (a) age; (b) county; (c) civil status; (d) profession

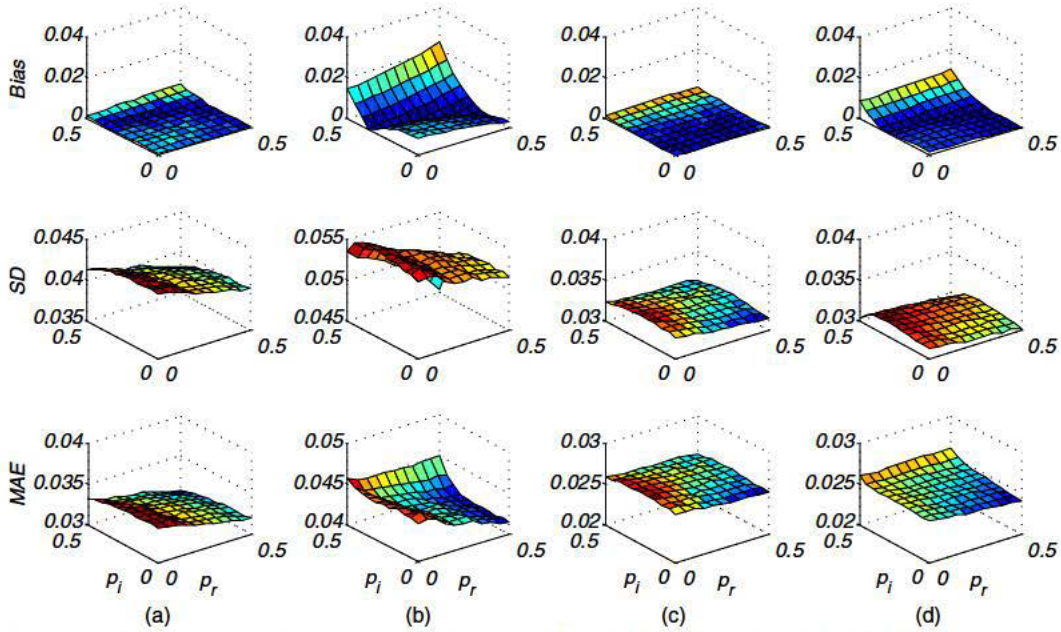


Fig. 10. RDS on $G1_{rand}$ with ignore and reject probabilities independent of the individuals' characteristics (simulations were repeated 100000 times for each combination of probabilities; number of seeds, 10; number of coupons, 3; sampling with replacement; seeds were randomly selected from the beginning of each simulation; estimates were calculated when sample sizes reached 500): (a) age; (b) county; (c) civil status; (d) profession

coupon generates one further participant and that each participant receives only one coupon. In our simulations, however, three coupons were used; see below. We modelled the effects of deviating from these ideal assumptions by letting each invited member have a probability of rejecting invitations. For each invited member, the probability of rejecting an invitation is called p_r. A rejected coupon was discarded and not reused for recruiting a new member.

In addition recruiters could potentially have difficulties remembering all their friends when considering whom to invite and when estimating the sizes of their personal networks. We modelled this by letting the recruiters ignore some of the edges when inviting members from his or her personal network. In the simulations, each edge was given a probability of being ignored. We call this probability p_i. An ignored edge has a zero probability of being selected and was not included in the network size of the participant when we calculated the RDSII estimates.

Additionally, we set the number of seeds to 10 and coupons to 3 to make sure that a sample could be recruited when p_r and p_i became large. Simulation results for G1, $G1_{add}$, and $G1_{rand}$ are displayed as surface plots in Figs 8–10.

When recruitment takes place with 'rejecting' and 'ignoring' in the original sparse network (Fig. 8), and in the $G1_{rand}$ networks (Figs 9 and 10), which are also sparse, the bias is small to moderate (up to 0.03). Simulation results on the edge-added networks reveal that, in these networks, changes in p_r and p_i do not affect the bias (Figs 9 and 10). Effects on the MAE are small in all networks (below 0.01).

In Fig. 8 it is also interesting to observe that, although increases in p_r and p_i increased the bias (most clearly for county and civil status), the MAE actually tended to decrease. The small differences that are seen when varying p_r and p_i, as well as the observed decreases in SD and

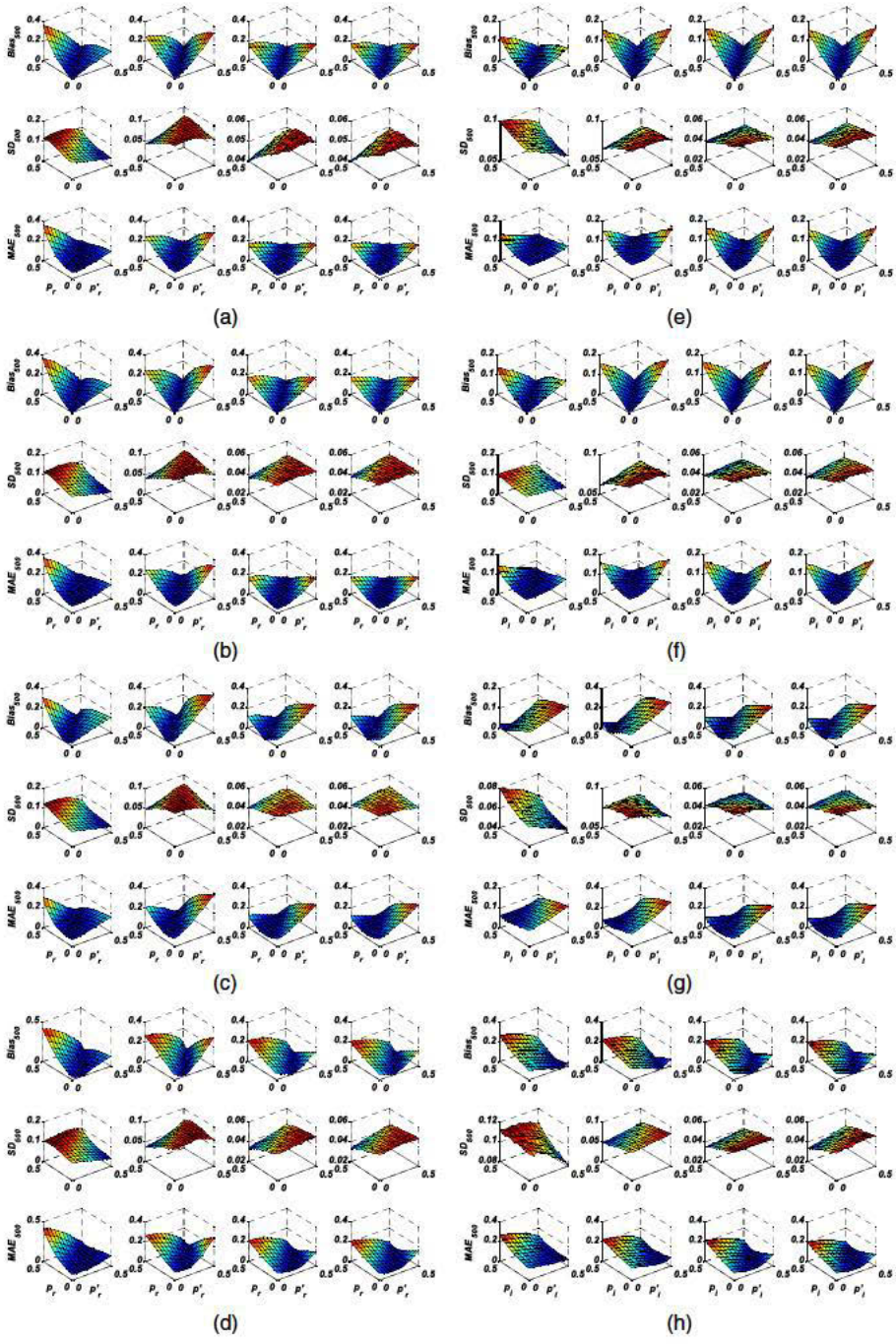


Fig. 11. RDS on G1 with ignore and reject probabilities dependent on the individuals' characteristics (simulations were repeated 10000 times for each combination; number of seeds, 10; number of coupons, 3; sampling with replacement; sample size, 500; seeds were randomly selected at the beginning of each simulation): (a) $p_i = 0, p'_i = 0$; (b) $p_i = 0.2, p'_i = 0.2$; (c) $p_i = 0.1, p'_i = 0.3$; (d) $p_i = 0.3, p'_i = 0.1$; from left to right in (a)–(d), age, county, civil status and profession; (e) $p_r = 0, p'_r = 0$; (f) $p_r = 0.2, p'_r = 0.2$; (g) $p_r = 0.1, p'_r = 0.3$; (h) $p_r = 0.3, p'_r = 0.1$; from left to right in (e)–(h), age, county, civil status and profession

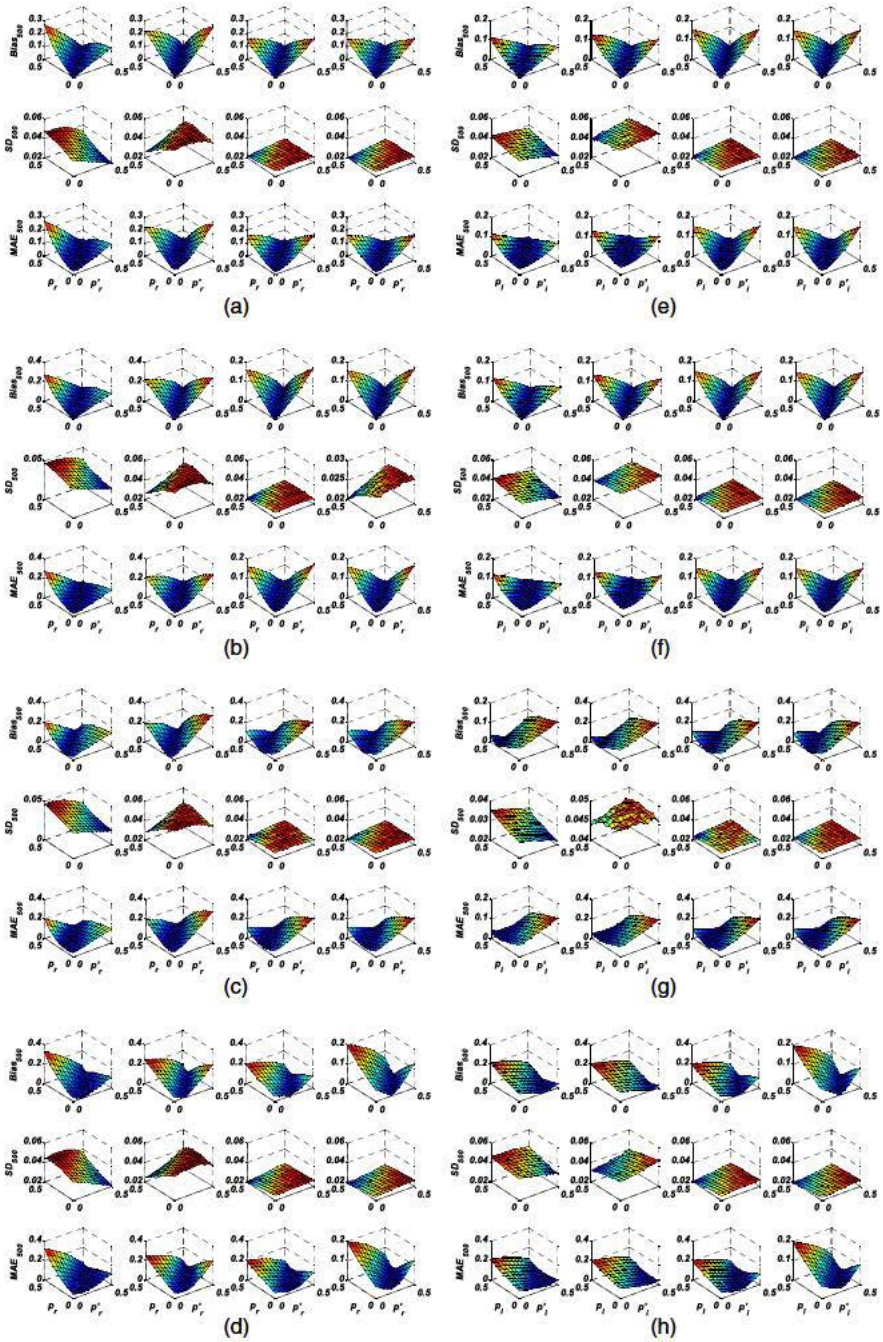


Fig. 12. RDS on $G1_{add}$ with ignore and reject probabilities dependent on the individuals' characteristics (simulations were repeated 10000 times for each combination of probabilities; number of seeds, 10; number of coupons, 3; sampling with replacement; seeds were randomly selected from the beginning of each simulation; estimates were calculated when sample sizes reached 500): (a) $p_l = 0, p'_l = 0$; (b) $p_l = 0.2, p'_l = 0.2$; (c) $p_l = 0.1, p'_l = 0.3$; (d) $p_l = 0.3, p'_l = 0.1$; from left to right in (a)–(d), age, county, civil status and profession; (e) $p_r = 0, p'_r = 0$; (f) $p_r = 0.2, p'_r = 0.2$; (g) $p_r = 0.1, p'_r = 0.3$; (h) $p_r = 0.3, p'_r = 0.1$; from left to right in (e)–(h), age, county, civil status and profession

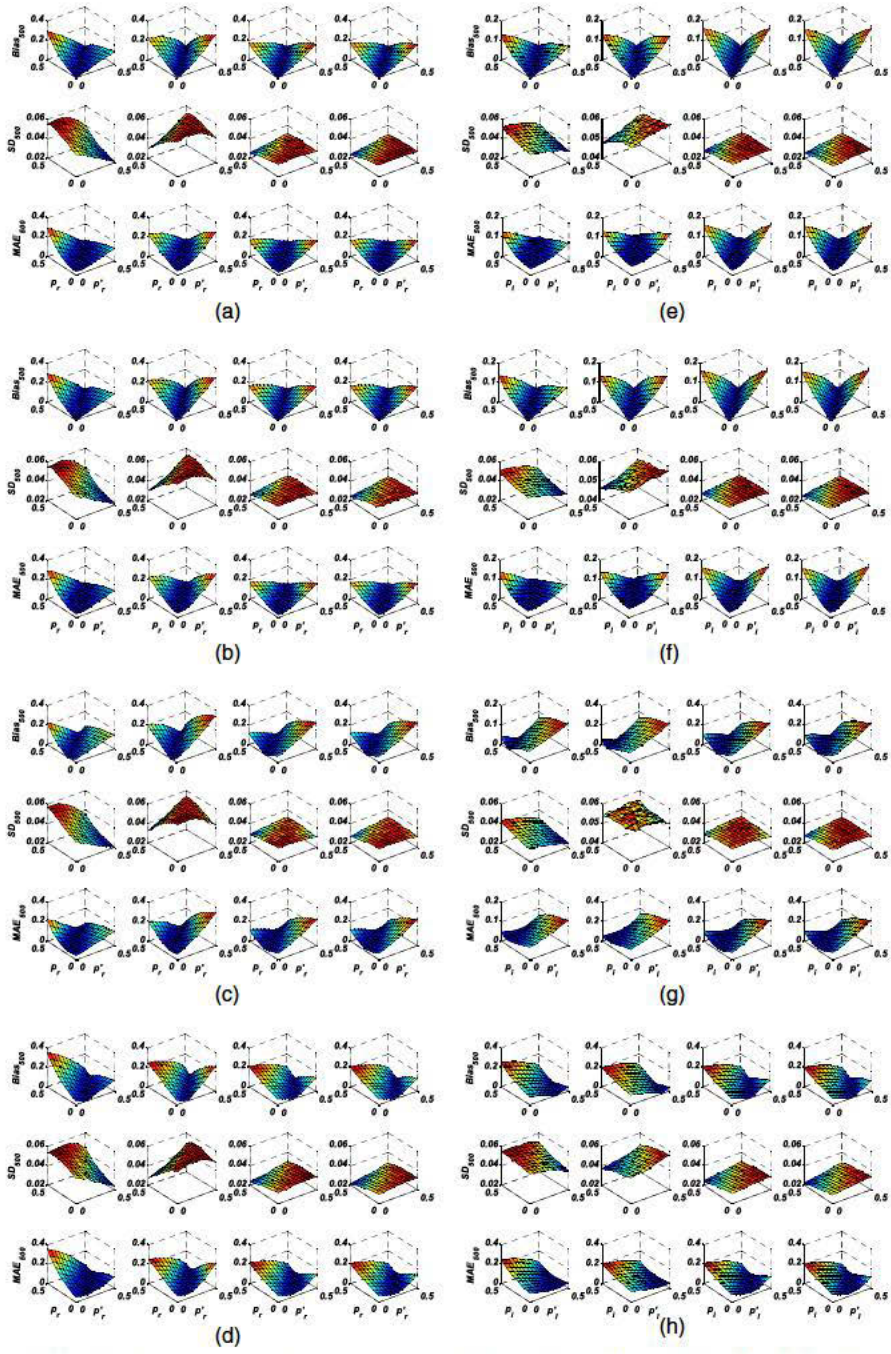


Fig. 13. RDS on $G1_{rand}$ with ignore and reject probabilities dependent on the individuals' characteristics (simulations were repeated 10000 times for each combination of probabilities; number of seeds, 10; number of coupons, 3; sampling with replacement; seeds were randomly selected from the beginning of each simulation; estimates were calculated when sample sizes reached 500): (a) $p_i = 0, p'_i = 0$; (b) $p_i = 0.2, p'_i = 0.2$; (c) $p_i = 0.1, p'_i = 0.3$; (d) $p_i = 0.3, p'_i = 0.1$; from left to right in (a)–(d), age, county, civil status and profession; (e) $p_r = 0, p'_r = 0$; (f) $p_r = 0.2, p'_r = 0.2$; (g) $p_r = 0.1, p'_r = 0.3$; (h) $p_r = 0.3, p'_r = 0.1$; from left to right in (e)–(h), age, county, civil status and profession

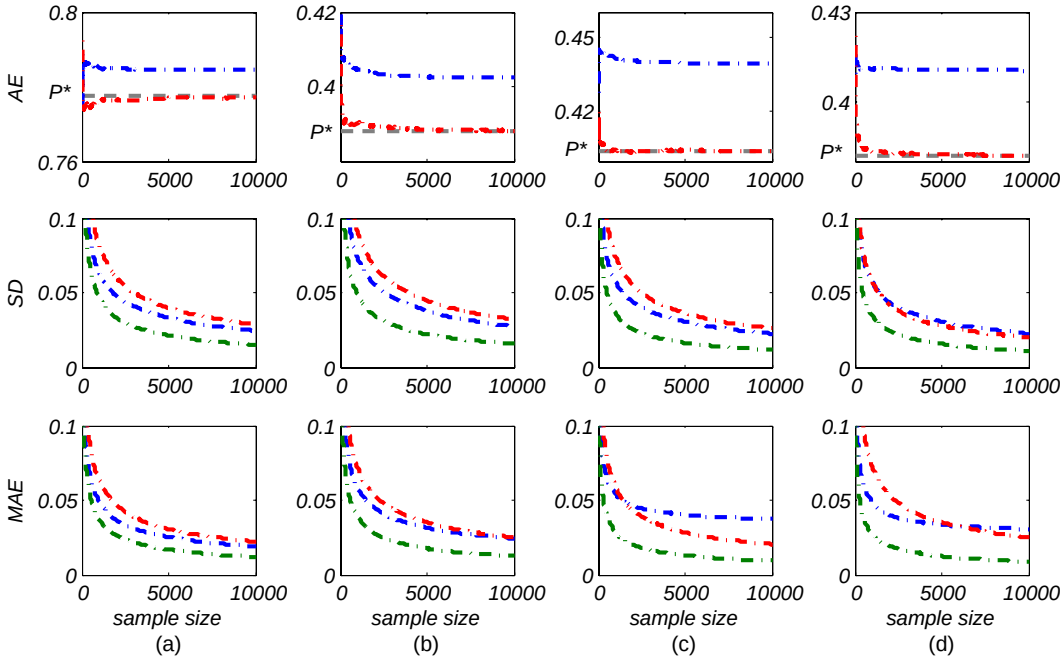


Fig. 14. RDS on $G1_{max}$ with preferential recruitment (number of seeds, 1; number of coupons, 1; sampling with replacement; ·-·-·, RDSII estimates; ·-·-·, eigenvector estimates; ·-·-·, RDSII estimates for recruitment with uniform probability; – – –, true population values; seeds were randomly selected at the beginning of each simulation): (a) age; (b) county; (c) civil status; (d) profession

MAE with increasing p_r and p_i, all imply a strong resistance of RDSII against these recruitment errors, as long as the recruitment chains can continue and generate the target sample size.

These simulations do not test all types of violation of assumptions (d) and (f). If we for example set p_i and p_r as dependent on any of our four outcome variables, errors could be much larger than described above. Indications of such non-random recruitment has been reported from a study of sex among men with men in Campinas, Brazil (de Mello *et al.*, 2008), and from a study of injecting drug users in Chicago (Scott, 2008). To evaluate the effects of non-random recruitment, we performed further simulations in which both the ignore and the reject probabilities differed, depending on group membership.

Let p_i be the probability that a member in the group of interest will be ignored by his friends when these friends are given the possibility to recruit, and let p_r be the probability that a coupon will be rejected by a member in the group of interest. Similarly, let p_i' and p_r' be the corresponding ignore and reject probabilities for members who are not in groups of interest. Surfaces for RDS with unequal recruiting probabilities are presented in Figs 11–13. We can see that, when the ignore or reject probabilities depended on the characteristics of the members, the RDS estimates gave large bias and error. Take the first simulation in Fig. 11, for example: when members who were born before 1980 rejected half of the invitations that were given to them and the members who were born after 1980 did not reject any invitations (p_i and p_i' both set to 0), the bias was over 0.3 for age.

When $p_i = p_i'$, the bias and MAE are small as long as $p_r = p_r'$ and vice versa. As both the ignore and the reject actions will reduce the inclusion probabilities of group members, they have similar effects on the RDS estimates and can compensate for each other. For example, when the

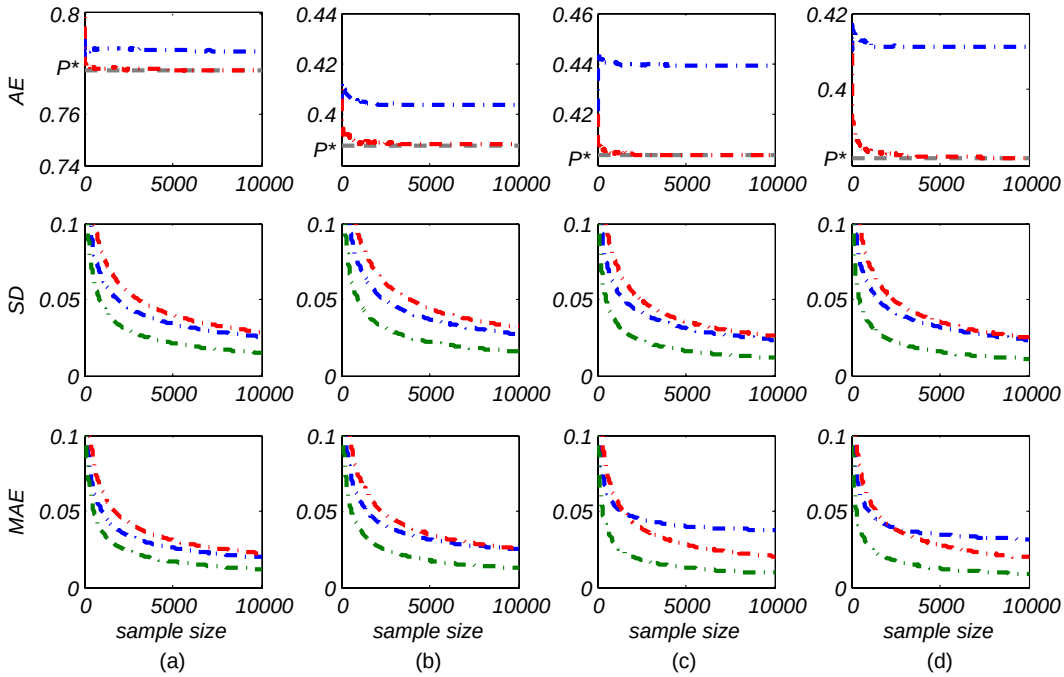


Fig. 15. RDS on $G1_{min}$ with preferential recruitment (number of seeds, 1; number of coupons, 1; sampling with replacement; ·-·-·, RDSII estimates; ·-·-·, eigenvector estimates; ·-·-·, RDSII estimates for recruitment with uniform probability; seeds were randomly selected at the beginning of each simulation): (a) age; (b) county; (c) civil status; (d) profession

fixed ignore probabilities were $p_i = 0.1$ and $p'_i = 0.3$, the minimum bias and MAE were when $p_r > p'_r$. For all the simulations, the values of the MAE were almost the same as those of the bias, revealing that, when the groups studied have different ignore or reject probabilities, the RDS estimates will virtually always be too high or too low. Although differences between groups in p_i can be compensated for by inverse differences between the groups in p_r, it can be hypothesized that such a combination of probabilities would be unusual in real life. As participants in RDS studies are rewarded for successful recruitments, a rational and self-interested participant would seek to ignore contacts whom he or she considers unlikely to accept an invitation. This would mean that groups with high p_r would also have a high p_i. Unfortunately, such combinations of p_i and p_r always give rise to the largest bias and MAE.

4.4. Preferential recruitment

There is empirical evidence of non-random recruitment in RDS studies. In a study of sex among men with men in Campinas City, Brazil (de Mello *et al.*, 2008), participants were reported most often to recruit close peers or peers who they believed practised risky behaviours. According to equation (2), it is easy to infer that RDSII would be biased when recruitment is a non-random selection among the edges of each node, as equation (4) is no longer the equilibrium.

A plausible non-random recruitment scenario would be that contacts with whom the recruiter interacts more frequently have a higher probability of being invited than those with whom the recruiter interacts only seldom. Simulation results for RDS on $G1_{max}$, in which respondents were supposed to recruit peers with a probability proportional to the edge weights, are presented in Fig. 14. We can see that the RDSII estimations were no longer unbiased for all groups. The

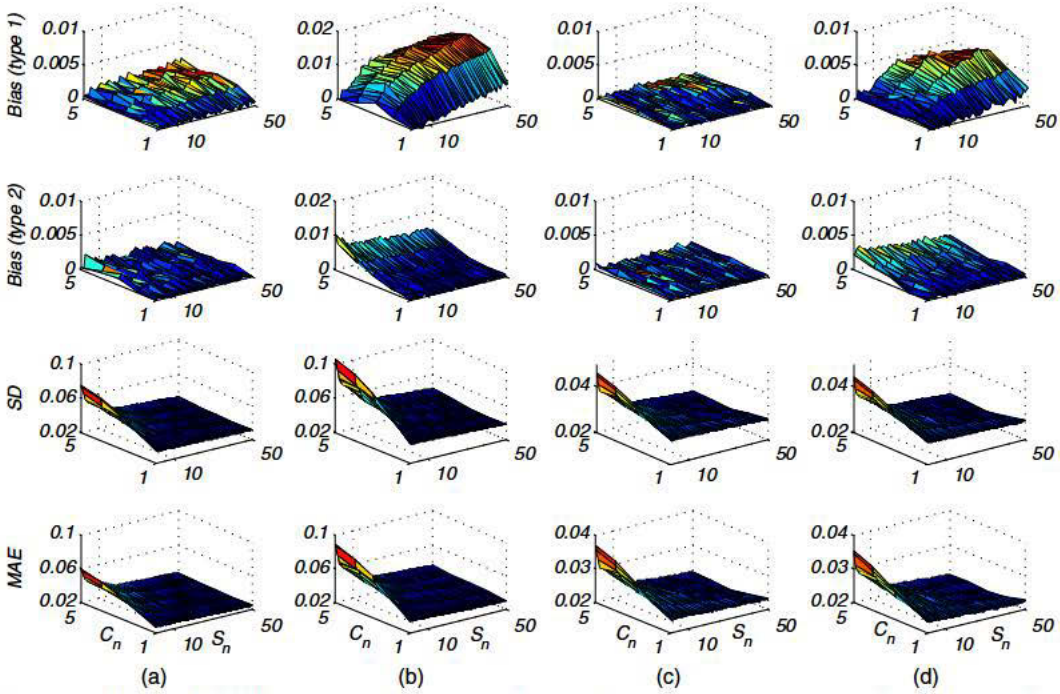


Fig. 16. Effects of RDSII estimates when varying the number of seeds and coupons within G1 (sampling was with replacement and the sample size was 500; the simulation was repeated 10000 times for each combination; C_n stands for the number of coupons and S_n for the number of seeds; seeds were not included in the sample): (a) age; (b) county; (c) civil status; (d) profession

biases were around 0.01, 0.02, 0.04 and 0.03 for the four groups. When we compare the SD and MAE of preferential recruitment with uniform recruitment, the preferential recruitment had larger SD and MAE values for all groups, indicating that, if the respondents prefer to distribute the coupons to their closer friends, larger RDSII estimation errors might occur. Again, when we use the eig estimations on the weighted network, they agree well with the true population.

Simulations on $G1_{min}$ reveal similar outcomes (Fig. 15).

4.5. Effect of seeds and coupons

Determination of the number of coupons per participant and the number and characteristics of the seeds are among the first problems that are encountered by researchers when preparing an RDS study. To evaluate the effects of variations in these parameters, we increased the number of coupons and seeds, with the seed(s) being selected either with uniform probability (type 1) or with probability proportional to the nodes' degree (type 2). Results are shown in Fig. 16. The difference between selection types for both SD and the MAE were minute, and we therefore do not show the SD and MAE for different selection types separately.

The first impression of this test is that an RDS study that is started by a type 2 selection seems to have a somewhat smaller bias than an RDS study that is started by a type 1 selection. The difference is slightly larger for age and county, which have larger homophilies. The number of seeds and coupons had small effects on the average RDSII estimates but had an obvious effect on the SD and MAE: both the SD and the MAE increased when the samplings used more coupons. This is probably because a study with a small number of coupons needs longer chains to

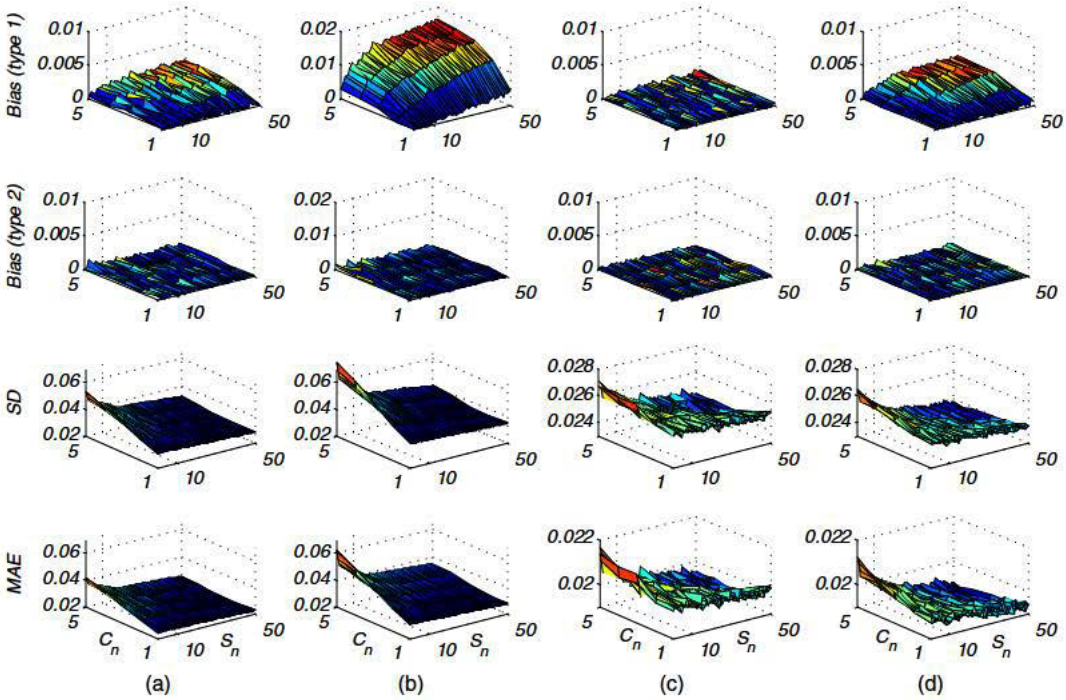


Fig. 17. Effects of varying the number of seeds and coupons in $G1_{add}$ when the sample size was 500, with replacement (the simulation was repeated 10000 times for each combination; C_n stands for the number of coupons and S_n for the number of seeds; seeds are not included in the sample): (a) age; (b) county; (c) civil status; (d) profession

reach the same sample size. Longer chains, in turn, are more likely to break out of homogeneous subnetworks and therefore become more representative of the overall population. Simulations on $G1_{add}$ and $G1_{rand}$ point to the same conclusions and are shown in Figs 17 and 18. Real life RDS studies cannot, however, use too few coupons or seeds as they will fail in generating sufficiently long recruitment chains.

5. Conclusions

Real social networks and the recruiting behaviour of people in those networks can barely meet all the theoretical assumptions underlying the RDS estimators. Therefore it is crucial to know how much estimates are affected by deviations from these assumptions. This paper describes, to the best of our knowledge, the first study simulating RDS studies within an actual hidden population and which in doing so can compare true population values with estimates obtained under various deviations from the RDSII assumptions.

Our simulations show that, when all the assumptions underlying RDSII are fulfilled, the results are excellent and asymptotically unbiased.

We show that the estimates that are adjusted by using weights based on the eigenvector of eigenvalue 1 for the transition matrix are always unbiased for all groups and networks. This is further the reason why RDSII is biased when the network is directed or when recruitment is a non-random selection from the participants' social networks. We should note that currently this eigenvector cannot be inferred from the empirical data in an RDS study where only information

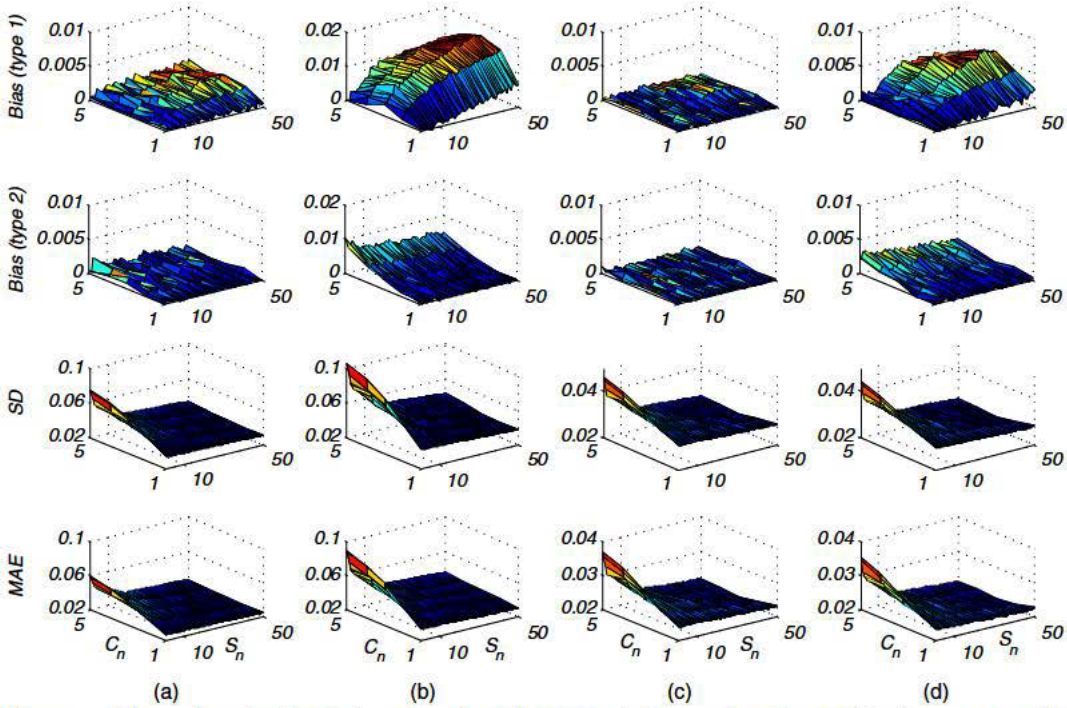


Fig. 18. Effects of varying the number of seeds and coupons in $G1_{rand}$ when the sample size was 500, with replacement (the simulation was repeated 10000 times for each combination; C_n stands for the number of coupons and S_n for the number of seeds; seeds are not included in the sample): (a) age; (b) county; (c) civil status; (d) profession

on recruitment chains and personal network sizes are known. However, our findings show that given knowledge about this eigenvector it is possible by using RDS to make unbiased estimates even on directed networks.

This study is to the best of our knowledge the first study showing the potential effects of preferential recruitment of close contacts. The results show that, when recruitment choices are weighted by the frequency with which people have communicated with each other, the bias, SD, MAE and design effect all increase.

An important result, which is supported by earlier research, is that sampling without replacement, which is used in practice, seems to give a lower design effect, SD and MAE than sampling with replacement. We also show that the design effects are not necessarily stable across sample sizes. The MAE, SD and design effect also decrease with increasing density of the network. These results are important to consider in future work.

RDSII shows strong resistance to recruitment error when the probabilities that individuals will ignore contacts and reject invitations are independent of the individuals' characteristics. However, if these probabilities are dependent on the individuals' characteristics and if these characteristics are correlated with the outcome characteristic that one wishes to estimate, the bias and MAE can become very large. As participants in RDS studies are rewarded for successful recruitments, rational and self-interested participants could be expected to ignore contacts who are considered to be less likely to accept invitations. Simulations show that such a combination of a group having a high probability of rejecting invitations and a high probability of being ignored can give rise to very large bias and error. We suggest that RDS studies should

routinely compare participants' reported network composition with actual recruitments to provide further empirical evidence on this issue from a wide variety of contexts.

In addition to testing the violation of assumptions, we also analysed some of the effects of network structure and homophily on RDSII. The results are consistent with previous studies: sparse networks with skewed degree distributions had larger error and bias, and estimations of groups with small homophily performed better than estimations among groups with high homophily.

The deviations from the assumptions that we have simulated in this paper can be modelled in different ways, which affect the conclusions. We have opted for deviations that we consider relevant to RDS studies in different contexts, but the simulations still represent subjective choices and do not cover all situations that are relevant to real life RDS studies. Moreover, although we have tried to make results more generalizable by varying the properties of the original network, the network characteristics that are actually picked for simulations do not reflect all types of networks, which could impact on the interpretations of the results. In summary we have shown the effects of a large number of deviations from assumptions. We show that the bias, MAE, SD and design effects are all affected by a large number of parameters. In further work, it would be valuable to run simulations of RDS studies with some realistic combinations of violations taken from real life studies as well as running simulations on other complete real life networks from diverse settings.

Acknowledgements

We thank two reviewers and the Joint Editor for many constructive suggestions on improving the paper. This work is funded in part by the *Riksbankens jubileumsfond* (grant P2008-0674). BJK acknowledges the support from the National Research Foundation of Korea (grant 2010-0008758), and XL thanks the China Scholarship Council (grant 2008611091).

Appendix A: Descriptions of network formation

The original male-to-male network G_{origin} included 31 945 profiles that belonged to homosexual men, and from which at least one message had been sent to another member. This network contained 364 746 directed edges, where each edge represented the sending of at least one message from a sender to a recipient during the data collection period.

When we keep only reciprocal links in the original network, i.e. if a message was sent from member i to member j, a connection between i and j was retained only if j also had sent a message to i. We thus obtained an undirected version of the original network, $G_{reciprocal}$, in which any edge was the result of a pair of directed edges. To make the network mimic a real social network and to meet the assumptions of RDS, we kept only members of the GCC of $G_{reciprocal}$ for our study, i.e. 16 082 members, which is the size of the networks that were studied throughout the paper.

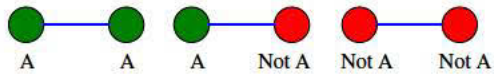
After defining the members to be studied in the network, we generated various networks by varying the inclusion criteria of edges. By keeping only the reciprocal edges between them, i.e. the GCC of $G_{reciprocal}$, we obtained an undirected network (G1) with an average degree of 6.74. By keeping both the reciprocal and the irreciprocal edges (the nodes defined by the GCC of $G_{reciprocal}$, but the edges retained as in G_{origin}), we obtain a directed network (G2) with an average degree of 17.2. Since all nodes belong to the GCC and the reciprocal edges were kept in both networks, all nodes would have the chance to be recruited with RDS sampling in either G1 or G2.

The generation processes of the other networks are as follows. (For simplicity, all edges discussed below are undirected.)

A.1. Generation of $G1_{add}$

For each property A, $G1_{add}$ is a network with the same nodes and homophily but with the average degree of the nodes increased by 20. The networks were created according to the following process.

For any edge, where each node either has or does not have the property A, there are three possible types of edge:



Let us denote the absolute numbers of edges of each type in G1 as n_1, n_2 and n_3, and the numbers of above types of link added are n_1', n_2' and n_3'. To increase the average degree by 20, the number of new links that we must add is $m' = n_1' + n_2' + n_3' = 16082 \times 20 = 321640$. To keep the homophily unchanged when these edges are added, we must keep the proportion of the three types of edges constant. The number of edges of each type that needs to be added is

$$n_k' = \frac{n_k}{\sum_1^3 n_i} m', \qquad \text{for } k = 1, 2, 3.$$

$G1_{add}$ is then obtained by repeating the procedure below.

(a) Randomly pick two nodes from the network.
(b) If there is no edge between the chosen nodes, determine the connections' configuration (type $k = 1, 2, 3$).
(c) If fewer than n_k' edges have already been added to the network, add a new edge of the type determined in step (b).
(d) Repeat the process until n_k' edges of each type have been added.

In such a way, the homophily of the network is unchanged and the density of the network is increased.

A.2. Generation of $G1_{rand}$

For each property A, $G1_{rand}$ is generated by rewiring existing edges in G1, so that the homophily with respect to property A remains unchanged. The generating processes are as follows.

(a) For each edge in G1, randomly pick one of the connected nodes to rewire.
(b) Randomly select a node from G1 with the same property as the node that was selected in step (a).
(c) Rewire the edge to the node selected in step (b).

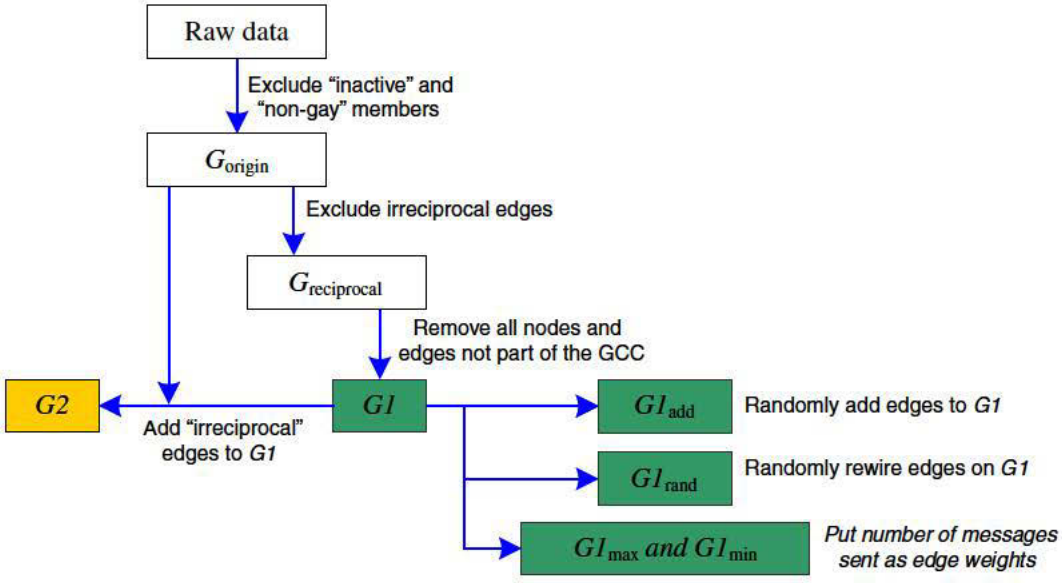


Fig. 19. Flow chart of the generation process of networks

(d) Repeat until all edges have been rewired at least once.
(e) If any node is not part of the GCC, repeat steps (a)–(c) until that does not happen.

In summary, the generation process of networks that was used in this paper can be depicted with the flow chart in Fig. 19, where the dark filled boxes are undirected networks and the light filled boxes are directed networks.

References

Abdul-Quader, A. S., Heckathorn, D. D., McKnight, C., Bramson, H., Nemeth, C., Sabin, K., Gallagher, K. and Des Jarlais, D. C. (2006) Effectiveness of respondent-driven sampling for recruiting drug users in New York City: findings from a pilot study. *J. Urb. Hlth*, **83**, 459–476.

Deaux E. and Callaghan, J. (1985) Key informant versus self-report estimates of health behavior. *Evaln Rev.*, **9**, 365–368.

Erickson, B. H. (1979) Some problems of inference from chain data. *Sociol. Methodol.*, **10**, 276–302.

Frost, S., Brouwer, K., Firestone Cruz, M., Ramos, R., Ramos, M. E., Lozada, R., Magis-Rodriguez, C. and Strathdee, S. (2006) Respondent-driven sampling of injection drug users in two U.S. Mexico border cities: recruitment dynamics and impact on estimates of hiv and syphilis prevalence. *J. Urb. Hlth*, **83**, 83–97.

Gile, K. J. and Handcock, M. S. (2010) Respondent-driven sampling: an assessment of current methodology. *Sociol. Methodol.*, **40**, 285–327.

Goel, S. and Salganik, M. J. (2009) Respondent-driven sampling as markov chain monte carlo. *Statist. Med.*, **28**, 2202–2229.

Goel, S. and Salganik, M. J. (2010) Assessing respondent-driven sampling. *Proc. Natn. Acad. Sci. USA*, **107**, 6743–6747.

Hastings, W. K. (1970) Monte carlo sampling methods using markov chains and their applications. *Biometrika*, **57**, 97–109.

Heckathorn, D. D. (1997) Respondent-driven sampling: a new approach to the study of hidden populations. *Socl Prob.*, **44**, 174–199.

Heckathorn, D. D. (2002) Respondent-driven sampling ii: deriving valid population estimates from chain-referral samples of hidden populations. *Socl Prob.*, **49**, 11–34.

Heckathorn, D. D. (2007) Extensions of respondent-driven sampling: analyzing continuous variables and controlling for differential recruitment. *Sociol. Methodol.*, **37**, 151–208.

Heckathorn, D. D. and Jeffri, J. (2001) Finding the beat: using respondent-driven sampling to study jazz musicians. *Poetics*, **28**, 307–329.

Heckathorn, D. D. and Jeffri, J. (2003) Social networks of jazz musicians. In *Changing the Beat: a Study of the Worklife of Jazz Musicians,* vol. III*, Respondent-driven Sampling: Survey Results by the Research Center for Arts and Culture*, pp. 48–61. Washington DC: National Endowment for the Arts Research Division.

Heimer, R. (2005) Critical issues and further questions about respondent-driven sampling: comment on Ramirez-Valles, et al. (2005). *Aids Behav.*, **9**, 403–408.

Johnston, L. G., Malekinejad, M., Kendall, C., Iuppa, I. M. and Rutherford, G. W. (2008) Implementation challenges to using respondent-driven sampling methodology for hiv biological and behavioral surveillance: field experiences in international settings. *Aids Behav.*, **12**, suppl., S131–S141.

Magnani, R., Sabin, K., Saidel, T. and Heckathorn, D. (2005) Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS*, **19**, suppl., S67–S72.

Malekinejad, M., Johnston, L. G., Kendall, C., Kerr, L., Rifkin, M. R. and Rutherford, G. W. (2008) Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic: review. *Aids Behav.*, **12**, suppl., S105–S130.

Marsden, P. V. (2005) Recent developments in network measurement. In *Models and Methods in Social Network Analysis* (eds P. J. Carrington, J. Scott and S. Wasserman), pp. 8–30. New York: Cambridge University Press.

McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001) Birds of a feather: homophily in social networks. *A. Rev. Sociol.*, **27**, 415–444.

de Mello, M., de Araujo Pinho, A., Chinaglia, M., Tun, W., Barbosa Júnior, A., Ilário, M. C. F. J., Reis, P., Salles, R. C. S., Westman, S. and Díaz, J. (2008) Assessment of risk factors for HIV infection among men who have sex with men in the metropolitan area of Campinas city, Brazil, using respondent-driven sampling. *Horizons Final Report*. Washington DC: Population Council.

Morris, M. and Kretzschmar, M. (1995) Concurrent partnerships and transmission dynamic in networks. *Socl Netwrks*, **17**, 299–318.

Page, L., Brin, S., Motwani, R. and Winogard, T. (1999) The pagerank citation ranking: bringing order to the web. *Technical Report 1999-66*. Stanford InfoLab, Stanford.

Ramirez-Valles, J., Heckathorn, D., Vazquez, R., Diaz, R. and Campbell, R. (2005) From networks to populations: the development and application of respondent-driven sampling among idus and latino gay men. *Aids Behav.*, **9**, 387–402.

Rapoport, A. (1980) A probabilistic approach to networks. *Socl Netwrks*, **2**, 1–18.
Rybski, D., Buldyrev, S. V., Havlin, S., Liljeros, F. and Makse, H. A. (2009) Scaling laws of human interaction activity. *Proc. Natn Acad. Sci. USA*, **106**, 12640–12645.
Salganik, M. J. (2006) Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *J. Urb. Hlth*, **83**, suppl. 1, 98–112.
Salganik, M. J. and Heckathorn, D. D. (2004) Sampling and estimation in hidden populations using respondent-driven sampling. *Sociol. Methodol.*, **34**, 193–239.
Schwarte, N., Cohen, R., Ben-Avraham, D., Barabasi, A. L. and Havlin, S. (2002) Percolation in directed scale-free networks. *Phys. Rev.* E, **66**, article 015104.
Scott, G. (2008) "They got their program and i got mine": a cautionary tale concerning the ethical implications of using respondent-driven sampling to study injection drug users. *Int. J. Drug Poly*, **19**, 42–51.
Volz, E. and Heckathorn, D. D. (2008) Probability based estimation theory for respondent driven sampling. *J. Off. Statist.*, **24**, 79–97.
Wang, J., Carlson, R. G., Falck, R. S., Siegal, H. A., Rahman, A. and Li, L. (2005) Respondent-driven sampling to recruit mdma users: a methodological assessment. *Drug Alc. Depend.*, **78**, 147–157.
Watters, J. K. and Biernacki, P. (1989) Targeted sampling: options for the study of hidden populations. *Socl Prob.*, **36**, 416–430.
Woess, W. (1994) Random-walks on infinite-graphs and groups: a survey on selected topics. *Bull. Lond. Math. Soc.*, **26**, 1–60.