Network Sampling: From Snowball and Multiplicity to Respondent-Driven Sampling

Douglas D. Heckathorn and Christopher J. Cameron

Department of Sociology, Cornell University, Ithaca, New York 14853-7601;
email: ddh22@cornell.edu, cjc73@cornell.edu

ANNUAL
REVIEWS **Further**

Click here to view this article's online features:

• Download figures as PPT slides
• Navigate linked references
• Download citations
• Explore related articles
• Search keywords

Keywords

respondent-driven sampling, network sampling, hidden populations, Markov, RDS, link-tracing sampling, social networks

Abstract

Network sampling emerged as a set of methods for drawing statistically valid samples of hard-to-reach populations. The first form of network sampling, multiplicity sampling, involved asking respondents about events affecting those in their personal networks; it was subsequently applied to studies of homicide, HIV, and other topics, but its usefulness is limited to public events. Link-tracing designs employ a different approach to study hard-to-reach populations, using a set of respondents that expands in waves as each round of respondents recruit their peers. Link-tracing as applied to hidden populations, often described as snowball sampling, was initially considered a form of convenience sampling. This changed with the development of respondent-driven sampling (RDS), a widely used network sampling method in which the link-tracing design is adapted to provide the basis for statistical inference. The literature on RDS is large and rapidly expanding, involving contributions by numerous independent research groups employing data from dozens of different countries. Within this literature, many important research questions remain unresolved, including how best to choose among alternative RDS estimators, how to refine existing estimators to make them less dependent on assumptions that are sometimes counterfactual, and perhaps the greatest unresolved issue, how best to calculate the variability of the estimates.

INTRODUCTION

Network sampling emerged owing to a recognition of the limitations of traditional methods such as household surveys. Some populations have attributes that make them hard to reach or "hidden" because it is infeasible to construct a sampling frame, that is, a list of population members from which the sample can be drawn. This difficulty arises when the target population is small relative to the general population and geographically dispersed, or when population membership involves stigma or the group has networks that are difficult for outsiders to penetrate (Sudman & Kalton 1986, Watters & Biernacki 1989, Spreen 1992, Brown et al. 1999). Groups with these characteristics are relevant to research in many areas, including public health (e.g., drug users and commercial sex workers), public policy (e.g., nondocumented immigrants and the homeless), and arts and culture (e.g., jazz musicians and aging artists). In developing countries, inadequate public records compound sampling problems, and consequently much of the general population qualifies, in sampling terms, as hidden.

Sampling hidden populations has traditionally involved a dilemma. Some studies have employed non-network-based probability sampling methods that provide incomplete coverage of the target population. For example, venue-based sampling (see, e.g., MacKellar et al. 1996) misses those who shun the large public venues, such as street corners and markets, from which subjects are recruited. Ramirez-Valles et al. (2005) show that the proportion missed can be substantial. Similarly, institutional sampling excludes respondents who are able to avoid incarceration in involuntary institutions such as jails and prisons, and those who shun voluntary associations such as support groups.

Other studies have employed network-based methods that provide more comprehensive coverage of the target population but yield only a convenience (i.e., nonstatistically valid) sample. The best known example is what is generally termed snowball sampling, in which research begins with a convenience sample of initial subjects who serve as seeds. Sampling then proceeds through network linkages, first from the seeds to the first wave, then from the first to the second wave, and so forth as the sample expands from wave to wave in the manner of a snowball growing as it rolls down a hill. Sampling stops when the target sample size has been attained. The major limitation of this method is that neither the selection of the initial subjects nor the selection of the subsequent waves is random, so the result is a convenience sample—a sample that does not provide the basis for validly inferring from the sample to the population from which it was drawn.

In the following, we examine the origins of snowball sampling methods, including how they evolved from a method for studying social networks for known populations, that is, populations for which a sampling frame was available. Proposed by Goodman (1961) and Coleman (1958), this approach to snowball sampling provides a useful tool for network analysis, but it was not applicable to the study of hidden populations, which of course lack a sampling frame. To avoid confusion between snowball as a convenience sampling method and its use as a probability sampling method to study social networks, the term snowball has now been replaced by the term link-tracing. This emphasizes more clearly the manner in which these samples function—that is, by tracing the links in the social network.

Interest in link-tracing designs has been fueled by recognition of their power to access members of hidden populations. As the literature on small world networks asserts, even in a nation as large as the United States, every person is indirectly associated with every other person through an average of approximately six intermediaries (Dodds et al. 2003). Therefore, everyone in the country could hypothetically be reached by the sixth wave of a maximally expansive link-tracing sample. Efforts have been made to transform these methods into probability sampling methods in order to overcome their limitations (Frank 1979, Snijders 1992, Frank & Snijders 1994), resulting

in a relatively new class of probability sampling methods termed adaptive designs (Thompson & Frank 2000).

A second type of network sampling method, multiplicity sampling, draws on respondents' knowledge of their own networks. Respondents are asked about events among those in their personal networks (i.e., acquaintances, friends, and those closer than friends). Generally this involves relatively rare events, such as having been robbed or murdered, or having contracted a disease such as AIDS. Suitable weighting systems, then, provide the basis for estimating the prevalence of these conditions in the population to compensate for cases when an event may be reported by more than one respondent. A limitation of this method is that it is only applicable to public events, that is, events that are likely to be known to one's network peers.

Respondent-driven sampling (RDS), a form of network sampling that Heckathorn (1997) introduced, combines both multiplicity and link-tracing methods. It has become popular because it provides a means by which mathematical adjustments can be applied to a sample to compensate for biases resulting from the network structures that affect sampling, thereby yielding a form of probability sample. Verdery et al. (2015) documented the popularity of the RDS method, noting that the ISI Web of Science database tags 642 academic articles with RDS listed as a topic and that these papers had been cited 10,217 times by 4,897 unique articles. The authors further showed that a search of the National Institutes of Health (NIH) RePORTER database revealed that the NIH had awarded more than $180 million to 448 projects and subprojects with "respondent driven sampling" as a topic. As Verdery et al. (2015) conclude, much of the popularity of RDS owes to the fact that it is a cost effective and rapid means of sampling hard-to-reach populations, an issue that has received increased attention across the social and health sciences.

In this review, we examine network sampling, beginning with its origin in snowball and multiplicity sampling, its evolution into a variety of more complex designs, and its subsequent development into the growing number of RDS-based sampling approaches and estimators, which are now being developed by a large number of independent research groups. Our survey of work on RDS includes assessment of the strengths and weaknesses of competing methods, recent theoretic and methodological innovations, and the major unresolved issues on which future work should focus.

ESTIMATION FROM NETWORK SAMPLES

Multiplicity Sampling

Multiplicity sampling was developed by Sirken (1970) in the late 1960s for sampling rare populations. A multiplicity survey differs from a conventional survey because each case may appear more than once. The approach is straightforward. For example, a telephone directory may have multiple entries for the same household. Consequently, when households are sampled from such a directory, they must be weighted based on multiplicity; for example, those with three phones have a weight only one-third that of households with a single listed phone. This approach is useful for increasing the efficiency with which household surveys can estimate the prevalence of rare events. The respondent is asked not only whether a condition affects his or her household, but also whether it affects a specified group of other households, such as those of surviving children and siblings. In this way, information regarding the event becomes available from not only the household surveyed, but also other households to which it is connected. Multiplicity arises because an event can be reported from multiple sources. In this way, the ability to detect rare events is increased. This approach was used by Laumann et al. (1989, p. 1186) to estimate the prevalence of homicide and AIDS cases in the United States. It involved starting with a random sample and

asking respondents about persons in their primary acquaintance networks, defined as those they knew personally as kin, friends, neighbors, coworkers, and casual or incidental acquaintances. The reports on incidence of homicide replicated official statistics quite well, but reports on incidence of AIDS were divergent from those from public health agencies; for example, prevalence among whites was underestimated, whereas that among minorities was overestimated.

The multiplicity approach was extended to link-tracing samples by Rothbart et al. (1982). They proposed adding to the survey a question regarding the number of eligible respondents known to the respondent. The size of this network then provides the basis for a multiplicity adjustment, in which respondents are weighted by the reciprocal of their network sizes (i.e., their degrees). The intuition underlying this approach is that respondents with large networks will have a greater probability of inclusion, because more recruitment paths lead to them, and thus respondents must be weighted by the reciprocal of their network sizes; that is, for any individual i with a degree D^i, the weight is $1/D^i$. The multiplicity weight for any individual i, MW^i, can therefore be defined as

$$MW^i = \frac{1}{D^i}.$$

1.

A limitation of multiplicity sampling is that it applies only to factors that are public knowledge. This applies to events such as homicide, which are matters of public record and are also a focus of the media. In contrast, a factor such as HIV infection is frequently private. For example, some celebrities who died of AIDS kept their status secret until the end (e.g., the actor Rock Hudson, or the journalist Randy Shilts). As we show, multiplicity sampling plays a role in RDS by providing a means for estimating the relative sizes of population groups' personal networks.

LINK-TRACING DESIGNS

Origin

Link-tracing designs originated in the Bureau of Applied Social Research at Columbia (Lazarsfeld et al. 1944, Merton 1949) in the 1940s. Based on the recognition that traditional survey methods were not adequate as a means to study the relationship between opinion leaders and followers, individuals were asked to name the people who influence them. From these, a second wave of people were then interviewed (for extended discussion see Handcock & Gile 2011). This approach was further developed by Coleman (1958) and Goodman (1961) as a means for studying the structure of social networks. Starting with a random sample of seeds, the sample expanded wave by wave to reveal influence patterns among physicians, and other phenomena.

Link-Tracing as a Convenience Sampling Method

Several years after Coleman's and Goodman's development of link-tracing designs, a new application of these designs emerged as a nonprobability approach to study hard-to-reach or, equivalently, hidden populations. In a review article, Biernacki & Waldorf (1981) observed that beginning with Becker's (1963) study of marijuana smokers and Lindesmith's (1968) classic study of opiate users, link-tracing designs had become both a standard technique in qualitative research and a topic in standard textbooks on sociological research methods. In this literature, "chain-referral sampling" served as the general term for link-tracing designs (Thompson & Frank 2000), Klovdahl's (1989) random walk approach, and others.

Samples of hidden populations must begin with a convenience sample of initial subjects, because if a random sample could be drawn, the population would not qualify as hidden in the first place.

These initial subjects serve as seeds, through which wave one subjects are recruited; wave one subjects in turn recruit wave two subjects; and the sample subsequently expands wave by wave like a snowball grows in size as it rolls down a hill.

In these convenience samples of hidden populations, sample proportions are used to estimate terms such as HIV prevalence. A typical problem with papers on this topic is that the introduction explains the limits of inference from a convenience sample, whereas the results are presented objectively and the discussion section then analyzes differences among groups as though these comparisons were based on a probability sample.

This transformation of the link-tracing method did not go without comment. Spreen (1992, p. 41) noted that "the historical purpose that Coleman had in mind (1958) . . . to study social structure changed into a total [*sic*] different purpose of . . . locating members of a special population." Thus, the transition to a convenience sampling method fit the needs of scholars who focused not on network structure, but instead on efficient and reliable means for accessing hidden populations.

RESPONDENT-DRIVEN SAMPLING

The use of link-tracing methods in research on hidden populations created a widespread perception that the method is strictly a convenience sampling method, and hence a method that does not support statistical inference. Erickson's (1979) article on problems of inference from chain data from hidden populations expressed the common wisdom. As she states, the sample begins with a convenience sample with bias of unknown magnitude and unknown direction, and this bias is then compounded in unknown ways as the sample expands from wave to wave. Hence, as applied to hidden populations, link-tracing is inherently limited to convenience sampling.

The judgment that link-tracing is a convenience method was challenged in a series of papers leading to the development of a new link-tracing method, RDS. This method was initially developed as part of a NIH/National Institute on Drug Abuse–funded HIV prevention project in Connecticut (Heckathorn 1997), with further funding from the Centers for Disease Control and Prevention (CDC) (Heckathorn et al. 2002). The method evolved incrementally in a series of papers that expanded and strengthened it. Because a different form of the method was presented in each paper, the term RDS refers not to a single method, but to a series of methods that have as their common core an effort to convert link-tracing into a probability sampling method.

The meaning of the term RDS has now been further expanded to refer not to the works of any single research team, such as Heckathorn and associates, but rather to the general enterprise of converting link-tracing into a probability sampling method. This enterprise is now pursued by numerous research groups, including those of Goel & Salganik (2009), Gile & Handcock (2015), Szwarcwald et al. (2011), Lu (2013), and Rohe (2015).

DEVELOPMENT OF RDS

The initial RDS paper (Heckathorn 1997) employed a Markov modeling of the peer recruitment process to show that, contrary to the conventional wisdom, bias from the convenience sample of initial subjects from which the sample inevitably began was progressively attenuated as the sample expanded wave by wave. This model employed data from peer recruitments to estimate the probability of recruitment across groups. These probabilities were organized into a recruitment matrix, specifying the probability of members of each group (for example, Hispanics), recruiting members from their own and each other group (for example, Whites, Blacks, Hispanics, and Others). These probabilities served as the transition probabilities of the Markov model. The model showed that as the sample expanded wave by wave, it approached an equilibrium that was

independent of the starting point, that is, independent of the convenience sample of seeds from which it began. The implication was that this sampling method could potentially become reliable if the number of waves was sufficiently large—that is, any selection of seeds would ultimately produce the same equilibrium sample composition. This insight provided a means for resolving the problem described by Spreen (1992, p. 49): "the central question in the methodological discussion about sampling and analyzing hidden populations is, basically: 'How to draw a random (initial) sample.'"

The solution is based on the insight that it does not matter if the initial sample is random if the number of waves reaches a threshold value large enough to eliminate bias from the initial selection of seeds. Furthermore, the analysis showed that bias from the seeds is reduced at a geometric rather than an arithmetic rate, a feature that accelerates the reduction of bias.

Using the Markov equilibrium as the population estimator, the initial RDS paper (Heckathorn 1997) showed that an RDS sample became self-weighting if groups had equal homophily, where homophily was defined by either the in-group bias estimator in Coleman's (1958) relational analysis or, equivalently, the inbreeding bias estimator in Rapoport's (1979) biased network theory. Though independently developed, both approaches to defining homophily are based on the same idea that a lack of structure in a network, that is, a lack of homophily, occurs when affiliation patterns fit random mixing, such that affiliations are formed randomly, irrespective of respondent characteristics. Positive homophily reflects a bias toward in-group affiliation, and negative homophily reflects a bias toward out-group affiliation, as in the case of sexual affiliations among heterosexuals.

This analysis demonstrated that under certain conditions, population estimates derived from a chain-referral sample could become not merely reliable, but also valid. However, the paper did not show how an unbiased estimate could be derived for cases where homophily was unequal.

These analyses (Heckathorn 1997, p. 192) also revealed a significant constraint on the method:

> When inbreeding terms are very large, reflecting mutual isolation of the system's groups, the approach to equilibrium slows. That is, when the terms exceed 0.99, scores of waves of recruitment may be required for equilibrium to be approximated. Therefore, given the practical limitations that constrain the number of waves of recruitment, this means that equilibrium will be reached only when inbreeding is not extreme. The implication is that when the boundaries separating groups are virtually impassable, RDS should be used to draw samples from within such groups, and not across them, even should inbreeding terms prove to be equal.

Such a possibility was illustrated by Heckathorn (1997). The analysis focused on a pair of cities with more than 99% recruitment from within each city and its adjoining area. To render the analyses tractable, the sample was divided into two subsamples, one for each city.

A subsequent paper (Heckathorn 2002) introduced a new RDS population estimator. Drawing not only on the data from the recruitment matrix, but also from self-reported network sizes, the estimators compensated both for differences in homophily across groups, and for differences in the mean degree (i.e., personal network size) across groups. This was accomplished through what was termed the reciprocity model. The essential idea was that in RDS, respondents recruit acquaintances, friends, and relatives, so their relationships tend to be reciprocal. Therefore, the number of ties linking any two groups must be the same in both directions—for example, monogamous marriage is a reciprocal relationship, so for any two groups X and Y, the number of Xs married to Ys must equal the number of Ys married to Xs. From this elemental network property, proportional group sizes can be calculated based on two types of network information: the proportion of cross-cutting ties between the groups and the relative sizes of each group's networks. Drawing the former information from the recruitment matrix, and the latter from the respondents'

self-reported network size, the new estimator was calculated based on these two types of network information (see Heckathorn 2002, p. 22). Given that this controlled for the effects of differences in homophily and network size across groups, this estimator became validly applicable across the full range of RDS data sets.

Heckathorn (2002) also clarified the role of homophily in RDS analysis, demonstrating that standard error increases exponentially as homophily increases. This reinforces the necessity of subdividing samples at high-homophily breakpoints, lest the variability of estimates become so great that they fail to yield meaningful information.

A related issue concerns unobserved heterogeneity. As an illustration, consider a study site with two neighborhoods, each of which has a different dominant ethnic group. An analysis that only considered ethnicity, that is, in which the groups forming the transition matrix involved only different ethnic groups, could give the impression of an area with ethnic integration. However, if the analysis were then expanded to include both ethnicity and neighborhood as variables, the analysis would reveal the ethnic segregation. More generally, if two variables are correlated, analyzing a single variable in isolation would fail to reveal the pattern of association. Hence, correlated variables should be analyzed in combination. In the simplest case, this would involve cross-tabs.

An additional paper (Salganik & Heckathorn 2004) derived the Salganik-Heckathorn RDS estimator and introduced a proof that this RDS estimator is asymptotically unbiased when the assumptions of the method are met, that is, bias is only on the order of 1/[sample size], and thus bias is minor in samples of substantial size. The specification of the assumptions that must be satisfied for the method to yield asymptotically unbiased population estimates provided a theoretically grounded means for assessing bias in RDS samples, through tests to see if the assumptions were approximated. Six assumptions were required for the proof:

1. Respondents know one another as members of the target population. For hidden populations, this means that population membership must create contact patterns, such as the ties created when jazz musicians perform together or when drug users buy and use drugs together.
2. The network of the target population forms a single component. This assumption is plausible if the network was created through a contact pattern, if it has small-world properties, or if it fits a power-law distribution. It is not plausible in the case of mutually isolated hidden populations, such as most tax cheats or child abusers.
3. Sampling occurs with replacement. Therefore, the sampling fraction must be small enough for a sampling-with-replacement model to be appropriate.
4. Respondents can accurately report their personal network size, that is, the number of people (alters) they know who fit the requirements of the study (such as drug injectors or jazz musicians). The RDS population estimators are insensitive to order-preserving linear transformations to network reports (i.e., multiplying self-reported network sizes by a positive constant), so linear inflations or deflations have no biasing effect. However, even when subjects are asked about a highly specific subset of those they know, such as other jazz musicians, error is inevitable. Wejnert (2009) tested several approaches to measure personal network size and found that the standard behavioral measure of the number of unique alters with whom the respondent interacted within a set time frame was consistently among the best degree measurement approaches. RDS estimators are sensitive to the presence of very low-degree nodes, so it is advisable to Winsorize the degree distribution to bring in the extreme values and therefore minimize the impact of measurement error.
5. Respondents recruit randomly from their personal networks. Therefore, the mean network size of a group can be estimated using a multiplicity approach, that is, all else being equal,

respondents are weighted by the reciprocal of their personal networks. This assumption is more plausible when members of the target population have easy and nonthreatening access to the research sites.

6. Respondents recruit only a single recruit, so recruitment effectiveness is uniform across groups.

The first five assumptions provide guidance both on when RDS is a suitable method and on suitable research designs. The sixth assumption is frequently counterfactual, because it is common for some groups to recruit more effectively than others. Hence, as discussed below, this assumption has been the subject of much theoretic work to minimize or eliminate potential for biases.

VOLZ-HECKATHORN: RDS II

A further development of the RDS method occurred in a paper by Volz & Heckathorn (2008), who derived the Volz-Heckathorn RDS estimator based on network principles. The founding insight is that when recruitment patterns are uniform across groups, respondents are sampled in proportion to their degree (i.e., their network size), and hence can be weighted by the inverse of their degree. The result is an estimator that resembles the multiplicity approach, with some notable differences, namely that degree data from seeds is treated as missing, because they generally consist of a convenience sample rather than having been recruited by peers. To distinguish it from the original RDS estimator, the new estimator was termed "RDS II." Like the original RDS estimator, the new estimator relies on Markov chain sampling theory (Metropolis et al. 1953, Hastings 1970), and it is also based on the theory of sampling with unequal probabilities (Hansen & Hurwitz 1943, Cochran 1977). In essence, to estimate the proportional size of a group (Volz & Heckathorn 2008, p. 86, equation 9), one adjusts the proportion of the group found in the sample by the ratio of the mean degree of the population and the mean degree of the group. That is, where P_A is the estimated proportion of group A, N_A is the number of A in the sample, N is the total sample size, D is the mean degree of the sample, and D_A is the mean degree of group A,

$$P_A = \frac{N_A}{N} \bullet \frac{D}{D_A}. \qquad 2.$$

Hence, if a group's mean degree is equal to that of the population as a whole ($D_A = D$), the ratio of total degree to the group's degree is equal to one ($D/D_A = 1$), so the RDS II population estimate is equivalent to the sample proportion, that is, $P_A = N_A/N$.

ADAPTIVE SAMPLING WITH LINK-TRACING DESIGNS

Thompson & Collins (2002) described link-tracing approaches that employ adaptive sampling. In adaptive sampling, information observed during the sampling process is used to direct future sampling efforts (Thompson & Seber 1995). In the context of link-tracing designs, only respondents who meet certain criteria might be asked to recruit peers. The authors showed that estimators from adaptive cluster sampling are appropriate if link-tracing samples are allowed to grow until exhausted and when the initial seeds are selected with known probability. The field implementation envisioned by the authors involves two stages: The first is a probability sample drawn from a larger population to identify subjects that meet the criteria of interest. Each identified subject is then the seed for a cluster discovered through exhaustive link-tracing. Thompson & Frank (2000) presented maximum likelihood estimation methods for data collected via these link-tracing designs. This sampling approach could be a good fit for populations where the group of interest is both large enough to be discovered through a random initial survey and small enough (or fragmented enough)

that exhaustive link-tracing is feasible. In contrast, RDS is best suited for well-connected groups that are difficult to discover via random sampling approaches, so these are complementary methods and the choice between them should be guided by the properties of the population of interest.

GILE'S SUCCESSIVE SAMPLING–BASED ESTIMATOR

Gile (2011) proposed an extension of the RDS II estimator, which has the advantage of controlling for bias from the sampling-with-replacement assumption. Recall that both the Salganik-Heckathorn and Volz-Heckathorn RDS estimators are based on the assumption of sampling with replacement. This is equivalent to assuming that samples are drawn from an infinite population or, equivalently, that the fraction of the population sampled—the sampling fraction—is indefinitely small. Given that this assumption is invariably counterfactual, an estimator that eliminated this source of bias had evident value.

Gile's successive sampling (SS) estimator is based on sequential sampling, a well-known elaboration of multiplicity sampling (Wald 1947). To illustrate how it works, consider again the example of applying the multiplicity method to sampling names from a phone book. After the first name is selected, the size of the phone book has been reduced by one name. Furthermore, given the elevated probability that the initial name that was selected was a person with more than one phone, the mean degree of the phone book (i.e., the mean number of phones per name) would decline. Therefore, were a second name to be drawn without adjusting the selection probabilities, there would be a possibility of bias because the mean degree of respondents in the phone book would be incorrectly specified.

Sequential sampling solves this problem by recomputing the phone book's number of names and degree for each name each time a name is selected and hence removed from the book. In this way, the estimates drawn from the sample remain unbiased, irrespective of the sampling fraction. For example, when most of the names in the book have already been selected, the remaining names will tend to have smaller numbers of phones, and this will in no way bias any of the sequential sampling estimates.

Gile tested her estimator using RDS data collected in three populations of drug users and men who have sex with men (MSM). In the analyses, she assumed that the sampling fractions varied from 50% to 95% and found that the SS estimator successfully controlled for bias under these circumstances. A limitation of the method is that it requires knowledge of the size and degree distribution of the sample. This is not problematic when sampling from a known population such as a phone book. However, as applied to hidden populations, whose sizes are—by definition—unknown, the assumption is problematic. Gile conducted a sensitivity analysis showing that if the error in the estimate of population size is 20% or less, bias is minor. However, studies of hidden populations have sometimes shown that estimates by key informants can be less accurate. For example, in the first application of capture-recapture to estimate the size of an injection-drug-using population (Heckathorn et al. 2002), a key informant, the head of the local vice squad, underestimated the number of injectors in his town by a factor of 11, that is, the number of injection drug users was 11 times larger than his estimate. Similarly, in a study of MSM in Kampala, public health officials claimed that the MSM population was tiny, consisting primarily of young westernized men who catered to sexual tourists. However, an RDS study revealed that the MSM population was not only much larger than the number who had become westernized, but that gay and bisexual men in Kampala are overwhelmingly Ugandan nationals from all parts of society (Kajubi et al. 2008).

A second limitation of the sequential-sampling estimator concerns its relevance in many RDS studies. A recent analysis by Barash et al. (2016) showed that for the reasonably small sampling

fractions (i.e., 20% or less) that are typical of applications of RDS in the field, the bias produced by the without-replacement assumption is negligible, and for larger sampling fractions (i.e., 40% or less), the bias is a small contributor to the variance in population estimates. For example, in the CDC's National HIV Behavioral Surveillance Injection Drug User (NHBS-IDU) study, the sampling fraction for the 23 study sites had a median of 2.3% and a range of 0.6% to 8% (Lansky et al. 2009). Similarly, in the National Endowment for the Arts–funded study of jazz musicians, the sampling fractions were 0.8% and 1.6% in New York and San Francisco, respectively (Heckathorn & Jeffri 2003). These studies illustrate contexts in which an RDS estimator need not control for bias from the sampling-with-replacement assumption.

GILE & HANDCOCK'S MODEL-ASSISTED ESTIMATOR

Gile & Handcock (2015) introduced an estimator that controls for dependence on the seed selection process when sample chain lengths are not sufficiently long or homophily is too severe to attenuate the influence of the seed on sample composition. The model-assisted method employs an exponential random graph model (ERGM; Lusher et al. 2013) of the underlying networked population that incorporates node degree, attribute values, and homophily. The network model is conditioned on network properties estimated from the RDS sample, and the RDS point estimates are based on sampling weights estimated from the network model. Computing the network model requires estimating the sampling weights, so estimation proceeds iteratively, alternating between estimating the sampling weights and estimating the parameters of the network model. This estimation procedure is highly computationally intensive, so it is significantly slower than other approaches, and this makes large-scale evaluation of the model-assisted estimator cumbersome (Gile & Handcock 2015, Verdery et al. 2015).

DUAL-COMPONENT RDS

Though the Volz-Heckathorn paper was published in 2008, it was submitted for publication in 2004. A further refinement of the RDS estimator was developed several years later (Heckathorn 2007). It provides a means for controlling bias from differential recruitment, in which some groups recruit more effectively than others, and hence their distinctive pattern of recruitment is overrepresented in the sample. More formally, differential recruitment bias (DRB) arises when a group recruits more effectively (i.e., more copiously) than others, and this differential recruitment effectiveness (DRE) is combined with a differential recruitment pattern (DRP), such as a preference for in-group rather that out-group recruitment. Hence, the two conditions for DRB are a combination of DRE and DRP. For example, in a study of New York City jazz musicians, females tended to recruit one another and also recruited more peers, so in combination these factors caused them to be oversampled (Heckathorn & Jeffri 2003).

The control for DRB is accomplished by dividing the RDS sampling weight, W, into an individual-based component calculated from respondents' self-reported network sizes, termed the degree component (DC), and a group-based weight, the recruitment component (RC), calculated from the recruitment matrix based on each group's patterns of recruitment of their own and other groups. These are defined such that the product of the two components would equal the sampling weight—that is, for a group X,

$$W_X = DC_X \bullet RC_X \qquad \qquad 3.$$

One benefit of this model is that it eliminates the need for the sixth assumption described in Salganik & Heckathorn (2004), in which each recruiter had only a single recruit. A second

benefit is that it refines the estimation of groups' network sizes by controlling for DRB in degree estimation. This involves replacing the prior multiplicity approach to estimating degree with the combination of multiplicity and link-tracing approaches characteristic of the RDS method (see Heckathorn 2007, p. 181).

The dual-component model also reduces constraints on RDS research design, permitting multistage designs to more effectively sample low-density sectors of social networks (Heckathorn 2007, p. 188). In essence, one draws a normal RDS sample. Given the tendency for high-degree individuals to be oversampled, this will undersample those of small degree. Hence, less information will be available on low-density regions of the network. This can be corrected by conducting a second-stage RDS, in which seeds are drawn exclusively from respondents of lowest degree. To the extent that low-degree respondents tend to cluster, this will provide deeper information on low-density regions of the network.

LINKED EGO NETWORKS

Lu (2013) suggested an estimator based on egocentric network data collected via RDS recruitment. The linked ego network approach collects data from each respondent about each alter in their ego network. RDS provides a framework for estimating inclusion probabilities, and the linked ego network approach extends this computation to compute the inclusion probabilities of each recruit's alters. Instead of relying on just the observed recruitment patterns, the transition probabilities are computed using the ego network composition, using self-reported count or proportion data. Simulations using real-world networks demonstrate that the linked ego network estimates control for differential recruitment and is fairly robust to biased and underreporting alters. The method is less robust to mislabeling alters, especially when the larger groups are systematically mislabeled. The limitation of the linked ego network approach is the assumption that respondents can correctly report the status of their peers. Although various demographic and status attributes can be accurately reported, information about health or private behaviors many not be available through ego net surveys. When feasible, linked ego network estimators can be used concurrently with other RDS estimators.

SAMPLING VARIANCE ESTIMATION

Point estimates without confidence intervals are of little practical use, but relatively little work on RDS variance estimation has translated into useful tools for practitioners. Variance estimation remains an open area of research. There are two main approaches for RDS sampling variance estimation: bootstrap and analytic approaches.

In general, RDS variance estimation approaches are designed to account for sampling variance from some or all of the following factors: personal network size estimation, recruitment patterns, and seed composition. The impact of seed selection on the RDS point estimates is known to approach nil as the number of sample waves increases (Salganik & Heckathorn 2004). The contribution of personal network size estimates was explored by Wejnert (2009), and different approaches were found to produce highly convergent estimates. The contribution of degree weights on the RDS estimates has two properties that minimize the contribution of measurement error in degree to the overall error in the RDS estimates. First, the degree weights appear as proportions relative to the other group sizes, so that error that is proportional to degree has no impact on the estimates. Second, low-degree respondents make a larger contribution to the estimates than high-degree nodes, so measurement error that scales with the personal network size has a minimal impact on the estimates, as most of the variance occurs in the estimates that contribute

the least. In cases with a few low-degree respondents, bootstrap approaches may yield higher variance. Winsorizing the degree distribution guards against these cases. Patterns of recruitment are the main contribution to sampling variance, and the recruitment patterns themselves are more difficult to measure and represent. Variance from seed selection and degree estimation are generally included in bootstrap approaches. Nearly all variance estimation techniques focus on better capturing the contribution of recruitment patterns on sampling variability.

The bootstrap approach has been gradually refined since its introduction. Heckathorn (2002) introduced a bootstrap-based estimate of the standard error of the RDS estimates that reflected the contribution of seed selection and transition probabilities between groups on the variance of the RDS estimates. Salganik (2006) further developed the bootstrap approach, incorporating dependence in the type of people recruited by each group of interest. Whereas Heckathorn (2002) used the observed transition matrix directly, the Salganik (2006) bootstrap resampled from the set of respondents recruited by each group of interest. The Salganik (2006) bootstrap estimates are available as part of the RDSAT software (Volz et al. 2012).

Volz & Heckathorn (2008) derived an analytic variance estimate for the Volz-Heckathorn estimator that performs well in the random networks on which it is tested. Neely (2009) recognized that both the Volz-Heckathorn variance estimate and Salganik (2006) bootstrap approaches treat RDS as a first-order Markov model and argued that their essential difference lies in the computational approach, and that the bootstrap incorporates variance from seed selection.

Neely (2009, p. 91) described a Bayesian bootstrap approach that accounts for several additional sources of variability, including "uncertainty in the estimated Markov transition matrix, in the seed selection process and in the relationship between degrees and sampling probabilities."

Yamanis et al. (2013) further extended the bootstrap approach, incorporating observed branching structures. In contrast to prior bootstrap approaches, they held the seeds fixed and used the Volz-Heckathorn estimator on the bootstrapped samples.

Gile & Handcock (2015) proposed a parametric bootstrap based on their model-assisted estimation procedure that draws simulated RDS samples from random ERGM networks conditioned on parameters estimated from the RDS sample. The distribution of estimates from these simulated samples can be used to generate confidence intervals for the point estimates. The parametric bootstrap performed adequately on simulated networks with homophily and seed bias.

Verdery et al. (2015) suggested two modifications to the algebraic Volz-Heckathorn variance estimator. They incorporated a measure of distance between nodes based on their common ancestor in the recruitment tree and allowed for higher-order Markov dependencies. Though these modifications improved the performance of the algebraic estimators, the authors found that the Salganik bootstrap method outperformed the algebraic variance estimates.

Though variance estimators tend to perform well when evaluated on simulated networks, their performance on real social networks can be discouraging. Goel & Salganik (2010) found that the Salganik (2006) bootstrapped confidence intervals were too narrow for simulated RDS samples drawn from two real-world networks. They reported coverage rates of 42–72% for 95% confidence intervals.

The discrepancy between the performance on simulated and real-world networks is due to differences in network structure and the distribution of the attribute of interest. This point was clearly articulated by Verdery et al. (2015), who argued that the first-order Markov assumption encoded in the RDS variance estimators results in too-narrow confidence intervals when the first-order assumption is violated. RDS estimation treats the population as a collection of relatively homogenous groups. These groups can be defined by the combination of several factors, but variance estimation generally proceeds from the assumption that individuals within groups are similar in their recruitment behavior. Verdery et al. demonstrated that current variance estimation

techniques are unable to detect or account for the impact of heterogeneous recruitment patterns within groups. When the intergroup ties are concentrated among a few nodes, they create what has been described as a bottleneck in the graph, a structure that is presently undetectable by statistical means but could potentially be identified by ethnographic study of the population of interest. (However, for a means for detecting bottlenecks, see below.) Groups should be defined with great care to maximize the recruitment pattern homogeneity within groups.

There remain opportunities for significant contributions to the practice of RDS by developing insights from the latest work into usable tools, developing creative new variance estimation approaches, and performing careful systematic comparisons among the proposed methods. Poon et al. (2009) demonstrated how alternative encodings for representing recruitment patterns reveal dependencies that are difficult to detect using a transition matrix framework. Future work might be able to leverage these alternative representations to produce bootstraps conditioned on higher-order dependencies.

CHOOSING AMONG ALTERNATIVE RDS ESTIMATORS

A frequent focus of RDS analyses has been comparing alternative estimators. For example, in the first application of RDS to a nonhidden population, university students, Wejnert (2009) compared the Volz-Heckathorn degree-based estimator and the Salganik-Heckathorn model-based estimator. He found that the point estimates were highly convergent. Such was not the case for the bootstrap variance estimates. The degree-based estimator had narrower confidence intervals and hence was more statistically efficient.

The dual-component estimator provides a means for explaining these findings. First, in the variables selected for analyses, differential recruitment happened to be minor, so the recruitment component of the sampling weight was small. Hence, the difference between the point estimates of the Salganik-Heckathorn and the Volz-Heckathorn estimators was minor. Second, the degree-based estimator was calculated based on a smaller amount of data—that is, only the respondents' self-reported degrees. In contrast, the model-based estimator was calculated based not only on degree data but also from peer-recruitment data. Given that the model-based estimator was calculated using a larger amount of data, there was also a greater potential source of variability in the data: hence, the wider confidence intervals for the model-based estimator.

It might seem that the above finding would indicate that the Volz-Heckathorn estimator is superior to the model-based estimator. However, this is not the case, because when differential recruitment is substantial, the Volz-Heckathorn estimator has a bias that is controlled by the model-based version. This form of bias is by no means rare. For example, in many drug studies, HIV-positive drug users are homophilous—they tend to cluster. Furthermore, they tend to recruit more effectively than do HIV-negative users. Hence, the combined effect is that they tend to be oversampled. A Volz-Heckathorn estimator would be blind to this source of bias, whereas a model-based estimator would correctly identify and control for it.

In summation, the dual-component paper provided a means for contextualizing the Volz-Heckathorn estimator, for if the recruitment component has a neutral value (i.e., $RC_X = 1$), the sampling weight is determined only by the degree component. In that case, the dual component and Volz-Heckathorn component become equivalent. Furthermore, as noted by Wejnert (2009), using the Volz-Heckathorn estimator provides a means for increasing statistical efficiency. Alternatively, in the more general and common case where the recruitment component is nonneutral, using a Volz-Heckathorn estimator would introduce bias by ignoring the differential recruitment that yielded the nonneutral recruitment component. Hence, in the general case, where the structure of the data set is unknown, it is prudent to use the Heckathorn (2007) dual-component estimator.

EVALUATIONS, CRITIQUES, AND REFINEMENTS OF RDS

A growing literature has evaluated the available RDS estimators, identifying sources of bias, critiquing existing research designs, and proposing new refinements to address unresolved issues. Especially notable are the articles by Salganik and colleagues (Goel & Salganik 2009, 2010; Gile et al. 2015). As noted above, the original RDS article (Heckathorn 1997) included a data set drawn from two adjoining cities with only a tenuous network connection—less than 1% of network links. This produced an extremely high homophily breakpoint of more than 95% separating the cities—homophily so extreme that approximating equilibrium would have required more than seven dozen waves of recruitment. Such high-homophily breakpoints are now termed bottlenecks (Goel & Salganik 2009). Gile et al. (2015) provided a useful means to identify such bottlenecks, termed a bottleneck plot. In essence, it analyzes the mean composition of each recruitment tree as it expands from wave to wave. In the presence of a bottleneck, some of the trees will converge at quite distinct equilibrium points, indicating that they are becoming stuck in different communities. In contrast, in the absence of a bottleneck, all of the seeds will converge at approximately the same convergence point, indicating that each tree's subsample is drawn from the same community.

Identifying such bottlenecks is important because, as Goel & Salganik (2009) have emphasized, they can even distort estimates for other variables in the data set. For example, if a city consists of two mutually isolated neighborhoods, ignoring the bottleneck separating them can distort estimates of features of the city, such as differences in ethnic or gender composition, HIV prevalence, or other terms. For an example, see the analysis of Poon et al. (2009) of injection drug users in Tijuana, in which sampling occurred in two neighborhoods on opposite sides of the city.

This potential for bias reinforces the importance of dividing samples at bottlenecks so that valid subsamples can be drawn from the networks on each side of the barrier, as was done in the original RDS paper (Heckathorn 1997). Furthermore, if the relative sizes of the populations from which the subsamples are drawn are known, the estimates can be recombined using standard statistical procedures to yield a valid estimate for the system as a whole. For examples of this procedure, the reader is referred to the CDC's NHBS system (Lansky et al. 2009) and a study of lower-tier workers in New York, Chicago, and Los Angeles (Bernhardt et al. 2009).

Goel & Salganik (2009) also offered further refinement to RDS methodology, which may be more limited in value. One is to discard data from early waves in a manner similar to the use of a burn-in phase in Markov chain Monte Carlo. A limitation of this approach is that it improves the validity and reliability of RDS estimates only when homophily is very high—indeed, when it is so high that RDS may not be a viable sampling approach unless the sample can be subdivided to eliminate bottlenecks. However, when homophily is low or moderate, early-wave recruitments are no more or less valid than those from later waves, for in RDS analyses, the unit of analysis is the recruiter-recruit relationship, a relationship that is not fundamentally different whether the recruitment is by a seed or a high-wave respondent.

A second recommendation is to stringently limit branching, that is, restrict recruitment quotas so each respondent can recruit only a limited number of peers. Here, too, the usefulness of this recommendation depends on the system's level of homophily. Using a hypothetical data set in which each group had 90% in-group recruitment, Goel & Salganik (2009) found that changing the quota from one to a range of zero to four widened the estimates' confidence intervals so much that the effective sample size was reduced by more than 70%. Equivalently expressed, the design effect was more than tripled.

Another analysis of the effects of branching on the width of confidence intervals in a much lower-homophily system (i.e., design effect of 1.5) found an opposite result, that branching had no significant effect (Heckathorn 2002, p. 28). Hence, just as the recommendation to discard early

waves may be useful only in systems with high homophily, so too may be the recommendation to severely restrict branching. To be most useful, recommendations regarding research design should be applicable even to the low- and moderate-homophily systems in which RDS is a viable sampling method.

Gile & Handcock (2010) also offered insightful analyses of the strengths and weaknesses of the RDS method. Like Goel & Salganik (2009), they used artificial data sets, and to better approximate actual RDS studies, they chose parameters which reflected the CDC's NHBS study. An intriguing finding is that, contrary to the findings of Goel & Salganik (2010), under certain circumstances, discarding the first two waves can increase rather than reduce bias. First, this reduces the number of waves in the sample, and second, the group from which seeds were drawn is differentially sampled in early waves, so eliminating them yields a bias.

Not surprisingly, Gile & Handcock (2010) also showed that selecting seeds randomly yields less bias than beginning from homogeneous seeds, for the latter increases the number of waves required for the sampling process to approximate equilibrium.

One of their more consistent findings concerns the superior performance of the degree Volz-Heckathorn estimator relative to the model-based Salganik-Heckathorn estimator. This is an artifact of the manner in which recruitment in their constructed networks took place. As is common in RDS simulation studies, they used a single nonbranching recruitment chain. This has the effect of precluding differential recruitment, the distinctive bias that the Salganik-Heckathorn estimator is designed to address. To see why this occurs, consider a two-category case; for example, in Gile & Handcock (2010), one category is infected, I, and the other is noninfected, N. If sampling begins with one or more seeds and proceeds without branching, then for an I to recruit an N, an N has to previously have recruited an I. In the simplest case, where sampling begins from a single seed, the recruitments from I to N must differ from the recruitments from N to I by no more than one. In essence, recruitment merely alternates between the two groups. Furthermore, given that cross-group recruitments are approximately equal, and the ratio of in-group to out-group recruitments is fixed owing to the assumption that recruits are a random sample from the respondent's personal neighborhood, there can be no DRE. Hence, there can be no DRB. The implication is that, when viewed from the standpoint of the dual-component model, the recruitment component must have a neutral value, so the sampling weight depends exclusively on the degree component. In that case, the Volz-Heckathorn estimator will be convergent with the Salganik-Heckathorn model-based estimator. The extra data upon which the Salganik-Heckathorn model relies will serve merely as a potential source of noise in the estimates.

FURTHER REFINEMENTS OF RDS ESTIMATORS

A common criticism of RDS is that its validity depends on multiple assumptions that frequently do not hold in the field. A recent paper illustrates how reliance on counterfactual assumptions can be reduced by suitable adjustments in the RDS estimator (Shi et al. 2016). Let us consider two examples. The first is an estimator that controls for bias from what is termed nonparticipation, in which a respondent accepts a recruitment coupon but then decides not to participate in the study. Differences in participation can arise from a variety of factors, such as differences in distance to the interview site when means of transportation are limited (McCreesh et al. 2011), or features of the interview site that are perceived as threatening to some groups of respondents and reassuring to others (Heckathorn 2008). The solution to this problem depends on the development of a measure for the nonparticipation rate that can be incorporated into the estimator. A second example, which was also proposed in Shi et al. (2016), controls for bias when status differences induce asymmetries in the recruitment patterns; for example, high-status respondents may be

reluctant to accept recruitment from those of lower status, or some groups may be reluctant to be recruited by other groups (Tajfel & Turner 1979, Tajfel 1982). A measure of this asymmetry is then incorporated in the estimator. We anticipate that this approach can be employed to reduce bias resulting from violation of other assumptions in RDS models. As these examples illustrate, a model-based approach to RDS statistical inference, as in Salganik & Heckathorn (2004) and Heckathorn (2007), provides a flexible means to account for the behavioral biases in recruitment and hence reduces reliance on counterfactual assumptions. In essence, the approach is to quantify the deviation from the assumption and then modify the estimator equation to compensate for this difference.

DISCUSSION

Compared with other statistical research methods, some of which have been refined over the course of many decades, RDS is in an early stage. Consequently, many important research questions remain open. At present, there is no consensus regarding which of the alternative point estimators is best. Given that different estimators are based on alternative assumptions about the structure of the data set, and depend on different types of data, it is likely that no single estimator will be optimal in all circumstances. Delineating these circumstances in an analytically rigorous manner will require considerable further work.

The greatest open question regarding RDS concerns variance estimation. Several distinct methods have been proposed based on various forms of either bootstrap (Heckathorn 2002, Salganik 2006, Neely 2009, Gile & Handcock 2015) or analytic methods (Volz & Heckathorn 2008, Verdery et al. 2015). An important tool in researching the variance of RDS estimators is the use of artificial data sets. When using actual data sets drawn from a hidden population, it is not possible to know the validity and reliability of a particular variance estimate. In contrast, by using an artificially contracted network with known parameters, multiple simulated RDS resamples can be drawn. The variance in the point estimates calculated from the set of resamples then provides what might seem to be a gold standard for assessing variance in RDS samples. However, whether it can meet that standard depends on whether the network structure from which the subsamples are drawn corresponds in essential ways to the structure of the hidden populations to which RDS is generally applied. Nonetheless, the results are instructive, and at a minimum they provide the means to compare alternative means for calculating variance. For example, Goel & Salganik (2010) and Wejnert et al. (2012) found that bootstrap and analytic variance estimators tend to underestimate the true variance of RDS estimators. This suggests that more work needs to be done. We expect that RDS will motivate methodological innovations for many years to come. Although important open questions remain, RDS remains the most practical sampling approach for hidden populations.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Respondent-Driven Sampling, Department of Mathematics, Stockholm University and Antonio Sirianni for helpful comments and advice.

LITERATURE CITED

Barash V, Cameron C, Spiller M, Heckathorn DD. 2016. Respondent-driven sampling—testing assumptions: sampling with replacement. *J. Off. Stat.* 32(1):29–73

Becker HS. 1963. *Outsiders: Studies in the Sociology of Deviance*. New York: Macmillan

Bernhardt A, Milkman R, Theodore N, Heckathorn D, Auer M, et al. 2009. *Broken Laws, Unprotected Workers: Violations of Employment and Labor Laws in America's Cities*. New York: Russell Sage Found.

Biernacki P, Waldorf D. 1981. Snowball sampling: problems and techniques of chain referral sampling. *Sociol. Methods Res.* 10:141–63

Brown LD, Eaton ML, Freedman DA, Klein SP, Olshen RA, et al. 1999. *Statistical Controversies in Census 2000*. Tech. Rep. 537, Dep. of Stat., Univ. Calif., Berkeley

Cochran WG. 1977. *Sampling Techniques*. New York: Wiley

Coleman JS. 1958. Relational analysis: the study of social organizations with survey methods. *Hum. Organ.* 17:28–36

Dodds PS, Muhamad R, Watts DJ. 2003. An experimental study of search in global social networks. *Science* 301(5634):827–29

Erickson BH. 1979. Some problems of inference from chain data. *Sociol. Methodol.* 10(1):276–302

Frank O. 1979. Estimation of population totals by use of snowball samples. *Perspectives on Social Network Research*, ed. P Holland, S Leinhardt, pp. 319–47. New York: Academic Press

Frank O, Snijders T. 1994. Estimating the size of hidden populations using snowball sampling. *J. Off. Stat.* 10:53–67

Gile K. 2011. Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. *J. Am. Stat. Assoc.* 106:135–46

Gile K, Handcock MS. 2010. Respondent-driven sampling: an assessment of current methodology. *Sociol. Methodol.* 40:285–327

Gile K, Handcock MS. 2015. Network model-assisted inference from respondent-driven sampling data. *J. R. Stat. Soc. A* 178(3):619–39

Gile KJ, Johnston LG, Salganik MJ. 2015. Diagnostics for respondent-driven sampling. *J. R. Stat. Soc. A* 178(1):241–69

Goel S, Salganik MJ. 2009. Respondent-driven sampling as Markov chain Monte Carlo. *Stat. Med.* 28(17):2202–29

Goel S, Salganik MJ. 2010. Assessing respondent-driven sampling. *PNAS* 107:6743–47

Goodman LA. 1961. Snowball sampling. *Ann. Math. Stat.* 32:148–70

Handcock MS, Gile KJ. 2011. Comment: on the concept of snowball sampling. *Sociol. Methodol.* 41(1):367–71

Hansen MH, Hurwitz WN. 1943. On the theory of sampling from finite populations. *Ann. Math. Stat.* 14(4):333–62

Hastings WK. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97–109

Heckathorn DD. 1997. Respondent driven sampling: a new approach to the study of hidden samples. *Soc. Probl.* 44(2):174–99

Heckathorn DD. 2002. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Soc. Probl.* 49:11–34

Heckathorn DD. 2007. Extensions of respondent-driven sampling: analyzing continuous variables and controlling for differential recruitment. *Sociol. Methodol.* 37:151–207

Heckathorn DD. 2008. *Assumptions of RDS: analytic versus functional assumptions*. Presented at CDC Consult. Anal. Data Collect. Respond.-Driven Sampl., Atlanta, GA

Heckathorn DD, Jeffri J. 2003. Social Networks of Jazz Musicians. In *Changing the Beat: A Study of the Work Life of Jazz Musicians. Vol. III. Respondent-Driven Sampling*, pp. 48–61. Washington, DC: Natl. Endow. Arts

Heckathorn DD, Semaan S, Broadhead RS, Hughes JJ. 2002. Extensions of respondent-driven sampling: a new approach to the study of injection drug users aged 18–25. *AIDS Behav.* 13(1):55–67

Kajubi P, Kamya MR, Raymond HF, Chen S, Rutherford GW, et al. 2008. Gay and bisexual men in Kampala, Uganda. *AIDS Behav.* 12:492–504

Klovdahl AS. 1989. Urban social networks: some methodological problems and possibilities. In *The Small World*, ed. M Kochen, pp. 176–210. New Jersey: Norwood

Lansky A, Drake A, Pham HT. 2009. HIV-associated behaviors among injecting drug users—23 cities, United States, May 2005–February 2006. *Morb. Mortal. Weekly Rep.* 58:329–32

Laumann EO, Gagnon JH, Michaels S, Michael RT, Coleman JS. 1989. Monitoring the AIDS epidemic in the U.S.: a network approach. *Science* 244(4909):1186–89

Lazarsfeld PF, Berelson B, Gaudet H. 1944. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Duell, Sloan and Pierce

Lindesmith AR. 1968. *Addiction and Opiates*. Chicago: Aldine

Lu X. 2013. Linked ego networks: improving estimate reliability and validity with respondent-driven sampling. *Soc. Netw.* 35(4):669–85

Lusher D, Koskinen J, Robins G. 2013. *Exponential Random Graph Models for Social Networks*. Cambridge, UK: Cambridge Univ. Press

MacKellar D, Valleroy L, Karon J, Lemp G, Janssen R. 1996. The Young Men's Survey: methods for estimating HIV seroprevalence and risk factors among young men who have sex with men. *Public Health Rep.* 111(Suppl 1):138–44

McCreesh N, Johnson LG, Copas A, Sonnenberg P, Seeley J, et al. 2011. Evaluation of the role of location and distance in recruitment in respondent-driven sampling. *Int. J. Health Geogr.* 10(1):56

Merton RK. 1949. Patterns of influence: a study of interpersonal influence and communications behavior in a local community. In *Communications Research, 1948–1949*, ed. P Lazarsfeld, F Stanton, pp. 180–219. New York: Harper

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. 1953. The Monte Carlo method. *J. Chem. Phys.* 21:1087–92

Neely WW. 2009. *Statistical theory for respondent-driven sampling*. PhD Thesis, Univ. Wisconsin, Madison

Poon AFY, Brouwer KC, Strathdee SA, Firestone-Cruz M, Lozada RM, et al. 2009. Parsing social network survey data from hidden populations using stochastic context-free grammars. *PLOS ONE* 4(9):e6777

Ramirez-Valles J, Heckathorn DD, Vázquez R, Diaz RM, Campbell RT. 2005. From networks to populations: the development and application of respondent-driven sampling among IDUs and Latino gay men. *AIDS Behav.* 9(4):387–402

Rapoport A. 1979. A probabilistic approach to networks. *Soc. Netw.* 2(1):1–18

Rohe K. 2015. Network driven sampling; a critical threshold for design effects. arXiv:1505.05461 [math.ST]

Rothbart GS, Fine M, Sudman S. 1982. On finding and interviewing the needles in the haystack: the use of multiplicity sampling. *Public Opin. Q.* 46:408–21

Salganik MJ. 2006. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *J. Urban Health* 83:i98–112

Salganik MJ, Heckathorn DD. 2004. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociol. Methodol.* 34:193–239

Shi Y, Cameron CJ, Heckathorn DD. 2016. Model-based and design-based inference: reducing bias due to differential recruitment in respondent-driven sampling. *Sociol. Methods Res.* **https://doi.org/10.1177/0049124116672682**

Sirken MG. 1970. Household surveys with multiplicity. *J. Am. Stat. Assoc.* 65:257–66

Snijders T. 1992. Estimation on the basis of snowball samples: how to weight? *Bull. Méthodol. Sociol.* 36:59–70

Spreen M. 1992. Rare populations, hidden populations, and link-tracing designs: What and why? *Bull. Methodol. Sociol.* 36:34–58

Sudman S, Kalton G. 1986. New developments in the sampling of special populations. *Annu. Rev. Sociol.* 12:401–29

Szwarcwald CL, Borges de Souza PR Jr., Damacena GN, Barbosa A Jr., Kendall C. 2011. Analysis of data collected by RDS among sex workers in 10 Brazilian cities, 2009: estimation of the prevalence of HIV, variance, and design effect. *J. Acquir. Immune Defic. Syndr.* 57:S129–35

Tajfel H. 1982. *Social Identity and Intergroup Relations*. Cambridge, UK: Cambridge Univ. Press

Tajfel H, Turner J. 1979. An integrative theory of intergroup conflict. In *The Social Psychology of Intergroup Relations*, ed. WG Austin, S Worchel, pp. 33–47. Monterey, CA: Brooks-Cole

Thompson SK, Collins LM. 2002. Adaptive sampling in research on risk-related behaviors. *Drug Alcohol Depend*. 68:S57–67

Thompson SK, Frank O. 2000. Model-based estimation with link-tracing sampling designs. *Surv. Methodol.* 26:87–98

Thompson SK, Seber AFG. 1995. *Adaptive Sampling*. New York: Wiley

Verdery AM, Mouw T, Bauldry S, Mucha PJ. 2015. Network structure and biased variance estimation in respondent driven sampling. *PLOS ONE* 10(12):e0145296

Volz E, Heckathorn D. 2008. Probability based estimation theory for respondent driven sampling. *J. Off. Stat.* 24:79–97

Volz E, Wejnert C, Cameron C, Spiller M, Barash V, et al. 2012. Respondent-Driven Sampling Analysis Tool (RDSAT) Version 7.1. *Software package*. **http://respondentdrivensampling.org**

Wald A. 1947. *Sequential Analysis*. New York: Wiley

Watters JK, Biernacki P. 1989. Targeted sampling: options for the study of hidden populations. *Soc. Probl.* 36(4):416–30

Wejnert C. 2009. An empirical test of respondent-driven sampling: point estimates, variance, degree measures, and out-of-equilibrium data. *Sociol. Methodol.* 39(1):73–116

Wejnert C, Pham H, Krishna N, Le B, DiNenno E. 2012. Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States. *AIDS Behav.* 16(4):797–806

Yamanis TJ, Merli MG, Neely WW, Tian FF, Moody J, et al. 2013. An empirical analysis of the impact of recruitment patterns on RDS estimates among a socially ordered population of female sex workers in China. *Sociol. Methods Res.* 42(3):392–425

ANNUAL REVIEWS
Connect With Our Experts

New From Annual Reviews:

Annual Review of Criminology
criminol.annualreviews.org • Volume 1 • January 2018

Co-Editors: **Joan Petersilia**, *Stanford University* and **Robert J. Sampson**, *Harvard University*

The *Annual Review of Criminology* provides comprehensive reviews of significant developments in the multidisciplinary field of criminology, defined as the study of both the nature of criminal behavior and societal reactions to crime. International in scope, the journal examines variations in crime and punishment across time (e.g., why crime increases or decreases) and among individuals, communities, and societies (e.g., why certain individuals, groups, or nations are more likely than others to have high crime or victimization rates). The societal effects of crime and crime control, and why certain individuals or groups are more likely to be arrested, convicted, and sentenced to prison, will also be covered via topics relating to criminal justice agencies (e.g., police, courts, and corrections) and criminal law.

Contents

Formal Organizations

Political and Economic Sociology

Differentiation and Stratification

Individual and Society

Demography

Urban and Rural Community Sociology

Policy

Indexes

Errata

An online log of corrections to *Annual Review of Sociology* articles may be found at
http://www.annualreviews.org/errata/soc