A SIMULATION-BASED FRAMEWORK FOR ASSESSING THE FEASIBILITY OF RESPONDENT-DRIVEN SAMPLING FOR ESTIMATING CHARACTERISTICS IN POPULATIONS OF LESBIAN, GAY AND BISEXUAL OLDER ADULTS

Respondent-driven sampling (RDS) is a method for sampling from a target population by leveraging social connections. RDS is invaluable to the study of hard-to-reach populations. However, RDS is costly and can be infeasible. RDS is infeasible when RDS point estimators have small effective sample sizes (large design effects) or when RDS interval estimators have large confidence intervals relative to estimates obtained in previous studies or poor coverage. As a result, researchers need tools to assess whether or not estimation of certain characteristics of interest for specific populations is feasible in advance. In this paper, we develop a simulation-based framework for using pilot data - in the form of a convenience sample of aggregated, egocentric data and estimates of subpopulation sizes within the target population - to assess whether or not RDS is feasible for estimating characteristics of a target population. In doing so, we assume that more is known about egos than alters in the pilot data, which is often the case with aggregated, egocentric data in practice. We build on existing methods for estimating the structure of social networks from aggregated, egocentric sample data and estimates of subpopulation sizes within the target population. We apply this framework to assess the feasibility of estimating the proportion male, proportion bisexual, proportion depressed and proportion infected with HIV/AIDS within three spatially distinct target populations of older lesbian, gay and bisexual adults using pilot data from the Caring and Aging with Pride Study and the Gallup Daily Tracking Survey. We conclude that using an RDS sample of 300 subjects is infeasible for estimating the proportion male, but feasible for estimating the proportion bisexual, proportion depressed and proportion infected with HIV/AIDS in all three target populations.

1. Introduction.

1.1. *Motivation and Overview.* Respondent-driven sampling (RDS) is a promising method for sampling from "hidden" or "hard to reach" target populations for which traditional probability sampling methods are impossible or prohibitively expensive (Heckathorn, 1997; Gile and Handcock, 2010). RDS begins with a small number of seeds, usually a convenience sample of

1

subjects from the target population. Seeds are entered into the study and asked to distribute coupons, which offer compensation for participation in the study, to contacts who also belong to the target population. Those who redeem coupons are in turn asked to distribute coupons to contacts in the target population. This process continues until a sample of the desired size is obtained.

Given enough seeds and coupons and sufficiently compelling compensation for participation, researchers have found that RDS often yields samples of the desired size (Malekinejad et al., 2008). However, several studies find that estimators derived from RDS samples, henceforth called RDS estimators, can perform poorly by giving biased estimates, point estimates that are extremely variable or interval estimates with poor coverage (Goel and Salganik, 2010; Gile and Handcock, 2010; McCreesh et al., 2012; Rohe, 2015). Consequently, researchers interested in using RDS for new target populations or characteristics of interest are in need of new methods for assessing RDS feasibility in specific contexts in advance of committing resources to a costly RDS study (Johnston et al., 2010; Kogan et al., 2011). Unfortunately, assessing whether or not RDS estimators are likely to perform poorly *before* conducting an RDS study is very difficult because RDS estimator performance depends on populations' social network structures and coupon distribution and redemption behaviors in complex ways (Salganik, 2006; Gile and Handcock, 2010; Goel and Salganik, 2010; Tomas and Gile, 2011; Lu et al., 2012; Wejnert et al., 2012; Rohe, 2015). Moreover, RDS estimator performance depends not only on the structure and behavior of the target population but the distribution of the characteristic of interest in the population (Gile and Handcock, 2010; Rohe, 2015). Without knowledge of or possibly unrealistic assumptions on the network structure and coupon distribution and referral behaviors of the target population (Rohe, 2015; Li and Rohe, 2015), closed form expressions relating network structure and coupon distribution and redemption behavior to RDS point and interval estimator performance are not available. Accordingly, we introduce a simulation based framework for assessing RDS feasibility. This framework involves repeatedly simulating RDS on synthetic networked populations, with network structure and coupon distribution and referral behaviors estimated from pilot data on the same population. We then assess RDS feasibility by examining different measures of RDS estimator performance across simulations.

Our work is motivated by the study of lesbian, gay and bisexual (LGB) older adult populations from three spatially distinct regions. The vast majority of existing research on LGB older adults is based on convenience samples that are unlikely to be representative of the target population (Fredriksen-

Goldsen and Muraco, 2010). At the same time, estimating characteristics of LGB older adult populations is of increasing importance for policy. The size of LGB older adult populations is growing as the population ages and LGB older adults are more likely to experience poor physical and mental health outcomes relative to the general population. RDS presents a promising method for sampling from and estimating characteristics of LGB older populations (Fredriksen-Goldsen and Muraco, 2010; Zea, 2010), however it may not be feasible.

Our framework assumes that researchers have access to pilot data in the form of aggregated, egocentric sample data. This pilot data may be from a convenience sample, a previous RDS study of the same population, or another source. Aggregated, egocentric data is comprised of counts of sampled subjects' contacts, stratified by various characteristics and has been used to estimate population network structure. Often, sampled subjects are referred to as egos and their contacts are referred to as alters. This kind of network data is very coarse, but nonetheless contains information about the distribution of contacts per subject in the population, i.e. the degree distribution, and how egos' characteristics relate to the characteristics of their alters, i.e. mixing totals. When alter counts are relatively small and the characteristics relevant to the social network structure are observed for egos and alters, this kind of data has been used to estimate measures of the structure of the population network (Admiraal and Handcock, 2016). However, aggregated egocentric data often includes very large alter counts and is "asymmetric" in practice. More characteristics are observed for egos than alters in asymmetric data, e.g. gender identity of egos may be observed while alter counts are pooled over gender identities. This is problematic because gender identity is relevant to network structure; the method given by Admiraal and Handcock (2016) will not allow gender identity to be used in modeling measures of the population network structure if gender identity is not observed for alters. Accordingly, we extend the work of Admiraal and Handcock (2016) to develop a novel statistical approach for estimating unknown social network structures from asymmetric aggregated social network data with large alter counts.

The RDS feasibility assessment framework we introduce is related to but distinct from previous work by Merli et al. (2015), who use data from a completed RDS study, combined with auxiliary data from a venue-based sample of the same population to simulate synthetic networked populations and then simulate RDS on the synthetic networked populations to assess the performance of the RDS estimates computed from the RDS study data. Whereas the work of Merli et al. (2015) gives a framework for assessing the

performance of RDS estimates *after* the RDS study has been conducted using detailed dyad-level social network information from the same data that the RDS estimates are computed from, our work provides a framework for assessing the performance of RDS estimates *before* conducting an RDS study using coarse, aggregated social network data from separate pilot data comprised of a convenience sample from the same population. In addition, our RDS feasibility framework introduced in this paper can be easily extended to investigate additional practical considerations, e.g. seed composition and number of coupons.

The rest of the paper proceeds as follows. First, we define and justify criteria for RDS feasibility. We then introduce our model for estimating measures of the population social network structure. We describe how synthetic networks can be simulated from these estimates, and how we simulate RDS on the synthetic networks. Using our framework for assessing RDS feasibility, we conclude that using an RDS sample of 300 subjects is infeasible for estimating the proportion male in all three LGB older adult populations, but feasible for estimating the proportion bisexual, proportion depressed and proportion infected with HIV/AIDS.

1.2. *RDS Feasibility Framework.* We define an RDS study of a specific characteristic in a specific population to be feasible if, over S RDS simulations on synthetic networked populations, we find that at least one RDS point estimator has effective sample size (ESS) greater than τ_S and a corresponding RDS interval estimator that has confidence interval width (WI) less than τ_W and coverage (CVR) higher than τ_C. The values τ_S, τ_W and τ_C are prespecified thresholds specific to the characteristic and population. If $\hat{\mu}_i^{\mathrm{RDS}}$ and $(\hat{l}_i^{\mathrm{RDS}}, \hat{u}_i^{\mathrm{RDS}})$ are RDS point and corresponding interval estimates for a specific characteristic obtained from the i-th RDS simulation with RDS sample size n, if $\hat{\mu}_i^{\mathrm{SRS}}$ is an estimate of the same characteristic computed from a simple random sample of the same size and if μ is the true proportion of the characteristic of interest in the synthetic network, the three measures of feasibility are:

$$ESS = n \left(\frac{\sum_{i=1}^{S} \left(\hat{\mu}_i^{\mathrm{SRS}} - \bar{\hat{\mu}}^{\mathrm{SRS}} \right)^2}{\sum_{i=1}^{S} \left(\hat{\mu}_i^{\mathrm{RDS}} - \bar{\hat{\mu}}^{\mathrm{RDS}} \right)^2} \right), \quad WI = \frac{1}{n} \sum_{i=1}^{S} (\hat{u}_i^{\mathrm{RDS}} - \hat{l}_i^{\mathrm{RDS}}) \text{ and}$$

$$CVR = \frac{1}{n} \sum_{i=1}^{S} \mathbb{1}_{\left\{ \hat{l}_i^{\mathrm{RDS}} \leq \mu \leq \hat{u}_i^{\mathrm{RDS}} \right\}}.$$

Because the best performing RDS point and interval estimators vary by network structure and coupon distribution and redemption behavior (Tomas

and Gile, 2011; Li and Rohe, 2015), we compute these measures of RDS feasibility for several common point and interval estimators for RDS data. In this paper, we consider the sample mean, the Salganik-Heckathorn (SH) and the Volz-Heckathorn (VH) estimators and compute 95% confidence intervals using exact binomial confidence intervals for the sample mean, the bootstrap procedure introduced in Salganik (2006) for the SH and VH estimators. For RDS feasibility studies in general, we recommend considering the successive sampling (SS) estimator as well, which corrects for finite population bias when the RDS sample size is large relative to the target population sizes (Gile, 2011). We do not consider the SS estimator for the feasibility analyses performed in this paper because the RDS sample size is small relative to the target population sizes and accordingly, VH and SS estimators perform identically as is often the case in practice (Barash et al., 2016).

We note that ESS is inversely related to the design effect (DE), $ESS = n/DE$, where n is the size of the RDS sample. DE is often considered in RDS research as a measure of RDS estimator performance (Salganik, 2006; Goel and Salganik, 2010; Wejnert et al., 2012). In this paper, we describe RDS feasibility in terms of ESS as opposed to DE because we believe approximate number of independent sampled subjects given by ESS is more interpretable to practitioners than the ratio of variances given by DE. Specifically, the DE is a measure relative to a SRS while ESS is an absolute measure of the sample design. If the decision was between a SRS and RDS then DE, the relative measure, would be primary. However, a SRS is not feasible here and the choice is between RDS and not doing a survey at all. Accordingly, ESS should be primary in this setting.

We require that RDS interval estimator coverage exceed a prespecified threshold, as opposed to requiring that RDS interval estimators obtain nominal, i.e. 95%, coverage, because all RDS variance estimators are known to underestimate RDS estimator variability (Verdery et al., 2015). A realistic threshold for coverage can be obtained according to studies of RDS interval estimator performance in practice (Salganik, 2006; Wejnert et al., 2012).

Considering all three of these quantities, ESS, WI and CVR separately may seem redundant. When a probability sample can be obtained, the variability of an estimator is often available in closed form as a function of sampling probabilities and sample size (Lohr, 2010). For a probability sample, nominal coverage will be achieved on average as long as sampling probabilities are known. Also, ESS and WI will directly determine each other and convey the same information about estimator performance. When RDS is used, ESS and WI as defined here do not convey the same information about estimator performance. Whereas ESS is computed from observed

variability of an RDS point estimator over RDS simulations, WI is computed from RDS interval estimators that use estimated RDS point estimator variability computed for each RDS simulation separately. Also, as RDS interval estimators often fail to achieve nominal coverage, CVR offers additional information on RDS estimator performance beyond ESS and WI.

With respect to setting values of τ_S, τ_W and τ_C in general, we recommend $\tau_S = n/3$, according existing literature on design effects which suggests assuming design effect between 2 and 4 when choosing the RDS sample size n, and $\tau_C = 0.8$ (Salganik, 2006; Wejnert et al., 2012). In the absence of existing interval estimates of the characteristic(s) of interest in the target population from previous studies, $\tau_W = \infty$ is appropriate because in the absence of any existing estimates a very wide interval could still be informative.

2. Data. In this section, we summarize the pilot data used to assess RDS feasibility for estimating characteristics of LGB older adult populations. We use a convenience sample from the Caring and Aging with Pride Study (CAP) (Fredriksen-Goldsen et al., 2013) and subpopulation size estimates of LGB older adult populations by metropolitan statistical area (MSA) from the Gallup Daily Tracking Survey (Gates, 2013).

2.1. *Caring and Aging with Pride Study.* The CAP Study, 2010-2011, was conducted through a collaboration with 11 agencies serving older adults in 8 spatially distinct MSAs across the US (Fredriksen-Goldsen et al., 2013). The majority of the agencies offered services for LGB and transgender older adults. Agencies invited older adults, age 50 and above, for whom contact information was available, to participate in the survey. Overall, 2,560 participants met the CAP study criteria. For a more detailed description of the data collection procedures, see Fredriksen-Goldsen et al. (2013). In this paper, we exclude transgender subjects due to small sample counts and focus on the three MSAs with enough data to estimate unknown parameters of our model for social network structure introduced in Section 3. These three MSAs contain 394, 712 and 340 subjects, respectively, after excluding subjects with missing data. For a more detailed description of the CAP data preparation, see Section 1 of the supplement.

Social Network Structures. We estimate social network structures from aggregated, egocentric sample data, which comprise reported counts of contacts stratified by sexual identity and gender identity as shown in Figure 1.

We account for sexual identity and gender identity in our estimates of social network structures. Sexual identity was measured by asking subjects to identify themselves as gay, lesbian, bisexual, heterosexual or other. Gender

Fig 1: Social network questions in the CAP questionnaire.

How many different lesbian, gay, bisexual, transgender or straight people (such as your friends, family members, colleagues, neighbors, etc.) have you interacted with (including talked to, visited with, exchanged phone calls or emails with, etc.) in a typical month?	
	Age 50 and older
a. Gay men:	_____
b. Gay women/lesbians:	_____
c. Bisexual men and women:	_____
d. Transgender men and women:	_____
e. Heterosexual or straight men and women:	_____

identity was measured by asking subjects to identify their gender identity as female, male or other. Subjects reported their age in years. For consistency with the Gallup subpopulation size estimates, age was dichotomized as younger adults under 65 years of age and older adults 65 years of age and older.

Coupon Distribution and Redemption Behaviors. Coupon distribution and redemption behaviors were measured by responses to the yes or no questions "Are you willing to be contacted for future projects addressing the aging and health needs of the LGBT community?" and "If you were participating in a future project addressing the aging and health needs of the LGBT community, would you be willing to spread the word about the project?"

Characteristics of Interest. As described in Section 1, we are interested in estimating the following proportions: male, bisexual, depressed and infected with HIV/AIDS. We described measures of subjects' gender and sexual identity earlier in this section, because they are used to estimate social network structures. Subjects' current depressive symptomology was measured using the Center for Epidemiological Studies Depression Scale (CES-D), 10-item short form (Radloff, 1977). Scores for the 10 items were summed and dichotomized, with a summed score of 10 or more indicating depression (Andresen et al., 1994). Subjects' current HIV/AIDS infection status was measured by asking subjects to indicate whether or not they had ever been told by a doctor that they have HIV/AIDS.

2.2. *Gallup Daily Tracking Survey.* The 2012 Gallup Daily Tracking Survey includes about 1,000 responses to the question "Do you, personally, identify as lesbian, gay, bisexual or transgender?", collected daily from June 1 to September 30, 2012 (Gates, 2013).

Subpopulation Size Estimates. We estimate subpopulation sizes of LGBT older adults by MSA, age group (< 65 vs. ≥ 65) and gender identity by combining estimates of population sizes from the 2010 Census with estimates of the proportions LGBT from the 2012 Gallup Daily Tracking Survey. Because the CAP sample proportion transgender is $\leq 2\%$ in all three MSAs, we use the Gallup estimates of LGBT older adult population sizes as working estimates of LGB older adult population sizes.

3. Methodology. In this section, we describe a method for estimating unknown social network structures and coupon distribution and redemption behavior from the pilot data described in Section 2, simulating synthetic networked populations consistent with our estimates of social network structures and subpopulation totals and simulating RDS on these synthetic social networks using estimates of coupon distribution and referral behavior. Throughout, we describe a network as a collection of nodes (egos) connected by undirected edges. If we select a single ego from the network, we refer to the nodes it is connected to as its alters.

3.1. *Estimation of Social Network Structure.* Building on the work of Admiraal and Handcock (2016), we first define a model for the unobserved social network structure that depends on several unknown parameters, and then relate these parameters to the pilot data described in Section 2.

Consider a network of N egos belonging to the target population with characteristics indexed by $i = 1, \ldots, I$ and $j = 1, \ldots, J$ and known subpopulation sizes $\boldsymbol{N} = (N_{11}, \ldots, N_{IJ})$. We refer to an ego with characteristics i and j as an ego of type ij. The characteristics indexed by i and j reflect heterogeneity in both egos' degrees and alter characteristics. In our application to LGB older adults, i indexes gay ($i = \mathcal{G}$) or bisexual ($i = \mathcal{B}$) sexual identity and j indexes male ($j = \mathcal{M}$) or female ($j = \mathcal{W}$) gender identity. As in Admiraal and Handcock (2016), we assume that social network structure is characterized by degree distributions, the distribution of alters per ego, and mixing totals, the propensity of egos to have alters of various types.

We define the network's degree distribution, stratified by characteristics indexed by i and j, as $\boldsymbol{D}_{ij} = (D_{ij,0}, \ldots, D_{ij,K})$. $D_{ij,k}$ denotes the number of egos of type ij with degree k, where $k = 0, \ldots, K$ and K is the maximum degree for all egos in the network. In contrast to Admiraal and Handcock (2016), who define a model for sexual contact networks which have small maximum degree, we define a model for social networks which are likely to have broad and heavy tailed tailed degree distributions with large maximum degree. This prompts us to adopt a parametric form for the degree distribution. We assume that the probability an ego of type ij has degree k follows

Matrix of Mixing Totals by Sexual Identity and Gender

	$\mathcal{GW}$	$\mathcal{GM}$	$\mathcal{BM}$	$\mathcal{BW}$
$\mathcal{GW}$	$M_{\mathcal{GW},\mathcal{GW}}$	$M_{\mathcal{GW},\mathcal{GM}}$	$M_{\mathcal{GW},\mathcal{BM}}$	$M_{\mathcal{GW},\mathcal{BW}}$
$\mathcal{GM}$	$M_{\mathcal{GM},\mathcal{GW}}$	$M_{\mathcal{GM},\mathcal{GM}}$	$M_{\mathcal{GM},\mathcal{BM}}$	$M_{\mathcal{GM},\mathcal{BW}}$
$\mathcal{BM}$	$M_{\mathcal{BM},\mathcal{GW}}$	$M_{\mathcal{BM},\mathcal{GM}}$	$M_{\mathcal{BM},\mathcal{BM}}$	$M_{\mathcal{BM},\mathcal{BW}}$
$\mathcal{BW}$	$M_{\mathcal{BW},\mathcal{GW}}$	$M_{\mathcal{BW},\mathcal{GM}}$	$M_{\mathcal{BW},\mathcal{BM}}$	$M_{\mathcal{BW},\mathcal{BW}}$

a negative binomial distribution:

$$(1) \quad p_{\mathcal{D}}\left(\rho_{ij}, \pi_{ij}; k\right) = \begin{cases} \binom{k+\rho_{ij}-1}{k} \left(1 - \pi_{ij}\right)^{\rho_{ij}} \pi_{ij}^k & \text{for } k < K \\ 1 - \sum_{k=1}^{K-1} \binom{k+\rho_{ij}-1}{k} \left(1 - \pi_{ij}\right)^{\rho_{ij}} \pi_{ij}^k & \text{for } k = K \end{cases},$$

where $\rho_{ij} > 0$ and $0 < \pi_{ij} < 1$ are unknown parameters that depend on characteristics indexed by i and j. The maximum degree and K is treated as fixed and known for computational tractability as explained in greater detail in Section 3.2.

It follows that the expected degree distributions for the network are:

$$(2) \qquad \boldsymbol{D}_{ij} = (D_{ij,0}, \ldots, D_{ij,K})$$
$$= \left(p_{\mathcal{D}}\left(\rho_{ij}, \pi_{ij}; 0\right), \ldots, p_{\mathcal{D}}\left(\rho_{ij}, \pi_{ij}; K\right)\right) N_{ij}.$$

We define a network's matrix of mixing totals as an $(IJ) \times (IJ)$ matrix with entries $M_{ij,i'j'}$, which give the number of edges connecting egos of type ij to alters of type $i'j'$. Because we assume that the edges are undirected, the mixing matrix is symmetric, i.e. $M_{ij,i'j'} = M_{i'j',ij}$.

For our application to LGB older adult populations, Table 1 represents mixing totals by sexual identity and gender, where $M_{\mathcal{GW},\mathcal{GW}}$ gives the number of edges connecting lesbian egos to lesbian alters in the population and $M_{\mathcal{GW},\mathcal{GM}}$ gives the number of edges connecting lesbian egos to gay male alters in the population.

Recalling that the network's expected degree distributions and mixing totals are *not* observed in the pilot data, we relate these unknown quantities to unknown parameters that can be estimated from the pilot data. We parametrize each mixing totals as a function of the degree distribution parameters ρ_{ij} and π_{ij} and mixing parameters $\alpha_{ij,i'j'}$,

(3)
$$M_{ij,i'j'} = \frac{\left(\sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{ij}, \pi_{ij}; k\right) k N_{ij}\right) \left(\sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{i'j'}, \pi_{i'j'}; k\right) k N_{i'j'}\right)}{\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{ij}, \pi_{ij}; k\right) k N_{ij}} \alpha_{ij,i'j'},$$

where $\alpha_{ij,i'j'}$ gives the relative propensity of egos of type ij to have alters of type $i'j'$ as opposed to alters of other types. Because we exclusively consider undirected networks, we assume $\alpha_{ij,i'j'} = \alpha_{i'j',ij}$.

Given these definitions for degree distributions and mixing totals, we have the following constraints on the row and column totals:

$$(4) \qquad \sum_{k=0}^{K} k D_{ij,k} = \sum_{i'=1}^{I} \sum_{j'=1}^{J} M_{ij,i'j'}.$$

Because $D_{ij,k}$ and $M_{ij,i'j'}$ are known functions of the unknown parameters $\boldsymbol{\rho}$, $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$, and known subpopulation sizes $\boldsymbol{N}$ in equations (2) and (3), this is a constraint on the unknown parameters $\boldsymbol{\rho}$, $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$.

We relate the unknown parameters $\boldsymbol{\rho}$, $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$ to the sample data by defining a sampling model for egos represented in the pilot data and defining sample degree distributions and mixing totals. To distinguish the sample degree distributions and mixing totals from the degree distributions $\boldsymbol{D}$ and mixing totals $\boldsymbol{M}$ introduced earlier, we refer to $\boldsymbol{D}$ and $\boldsymbol{M}$ as population degree distributions and mixing totals and denote the corresponding sample quantities with lowercase letters.

Throughout this paper, we assume that the sampled egos in the pilot data are a simple random sample from the target population. We view this as a working assumption; we know that sampled egos in the pilot data are likely a convenience sample, however we argue that assuming they are a simple random sample is sufficient for estimating the approximate social network structure for the purpose of simulating RDS on synthetic social networks.

To reflect features of our pilot data on LGB older adults, we assume each sampled ego reports counts of alters stratified by alter characteristics indexed by i and j. We assume that i indexes characteristics that are observed for both egos and the alters they report, whereas j indexes characteristics that are observed for egos in the sample and alters with certain values the of characteristic indexed by i. In our pilot data on LGB older adults where i indexes sexual identity and j indexes gender identity, egos report counts of gay alters by gender identity and counts of bisexual alters pooled across gender identity.

Without loss of generality, we assume that the types are ordered such that alter characteristics indexed by j are observed when $i \leq \tilde{I}$ and not observed when $i > \tilde{I}$. We denote reported alter counts for the l-th sampled ego of type ij, which we assume are reported without error, by

$$\boldsymbol{c}_{ijl} = \left(c_{ijl,11}, \ldots, c_{ijl,\tilde{I}J}, c_{ijl,(\tilde{I}+1)\cdot}, \ldots, c_{ijl,I\cdot} \right),$$

where $c_{ijl,i'j'}$ and $c_{ijl,i'.} = \sum_{j'=1}^{J} c_{ijl,i'j'}$ are the l-th sampled ego of type ij's reported number of alters of type $i'j'$ and i', respectively, and $l = 1, \ldots, n_{ij}$. In our study of LGB older adults, we abuse notation and write $\tilde{I} = \mathcal{G}$ and assume $\mathcal{B} > \mathcal{G}$.

Using this notation, sample degree distributions are defined as

$$
(5) \qquad \boldsymbol{d}_{ij} = (d_{ij,0}, \ldots, d_{ij,K})
$$

$$
= \left(\sum_{l=1}^{n_{ij}} \mathbb{1}_{\left\{ \sum_{i'=1}^{\tilde{I}} \sum_{j'=1}^{J} c_{ijl,i'j'} + \sum_{i'=\tilde{I}+1}^{I} c_{ijl,i'.} = 0 \right\}}, \ldots, \right.
$$

$$
\left. \sum_{l=1}^{n_{ij}} \mathbb{1}_{\left\{ \sum_{i'=1}^{\tilde{I}} \sum_{j'=1}^{J} c_{ijl,i'j'} + \sum_{i'=\tilde{I}+1}^{I} c_{ijl,i'.} = K \right\}} \right).
$$

Under simple random sampling of egos, the sample degree distribution for egos of type ij, sexual identity i and gender identity j in our pilot data, follows a multinomial distribution with total n_{ij}:

$$
(6) \qquad \boldsymbol{d}_{ij} | n_{ij} \sim \mathcal{MN} \left(n_{ij}, (p_{\mathcal{D}} (\rho_{ij}, \pi_{ij}; 0), \ldots, p_{\mathcal{D}} (\rho_{ij}, \pi_{ij}; K)) \right).
$$

Elements of the $(IJ) \times \left(\tilde{I}J + \left(I - \tilde{I} \right) \right)$ matrix of sample mixing totals are defined as $m_{ij,i'j'} = \sum_{l=1}^{n_{ij}} c_{ijl,i'j'}$, for $i \leq \tilde{I}$, and $m_{ij,i'.} = \sum_{l=1}^{n_{ij}} c_{ijl,i'.}$, for $i > \tilde{I}$. The matrix of sample mixing totals for our application to LGB older adults is shown in Table 2, alongside the corresponding matrix of population mixing totals, where $M_{\mathcal{GW},\mathcal{B}.} = M_{\mathcal{GW},\mathcal{BM}} + M_{\mathcal{GW},\mathcal{BW}}$. Given these definitions for sample degree distributions and mixing totals, the following constraint on the row totals of the sample matrix of mixing totals holds:

$$
(7) \qquad \sum_{k=0}^{K} k d_{ij,k} = \sum_{i'=1}^{\tilde{I}} \sum_{j'=1}^{J} m_{ij,i'j'} + \sum_{i'=\tilde{I}+1}^{I} m_{ij,i'.}.
$$

The constraints given by equation (7) suggest a multinomial model for each row of the sample matrix of mixing totals, given the sample degree distributions. Under simple random sampling of egos, the multinomial probabilities will be functions of the unknown parameters $\boldsymbol{\rho}$, $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$ and known subpopulations sizes $\boldsymbol{N}$ derived from equation (3).

In contrast to Admiraal and Handcock (2016), this is a rectangular matrix of mixing totals. As a result, some unknown mixing parameters $\boldsymbol{\alpha}$ are not identifiable. Consequently, we reparametrize the cells of the matrix of population mixing totals before deriving the multinomial probabilities for

TABLE 2
Matrix of Mixing Totals by Gender and Sexual Identity, Collapsed

(a) Sample   (b) Population

	$\mathcal{GW}$	$\mathcal{GM}$	$\mathcal{B}$		$\mathcal{GW}$	$\mathcal{GM}$	$\mathcal{B}$
$\mathcal{GW}$	$m_{\mathcal{GW},\mathcal{GW}}$	$m_{\mathcal{GW},\mathcal{GM}}$	$m_{\mathcal{GW},\mathcal{B}\cdot}$	$\mathcal{GW}$	$M_{\mathcal{GW},\mathcal{GW}}$	$M_{\mathcal{GW},\mathcal{GM}}$	$M_{\mathcal{GW},\mathcal{B}\cdot}$
$\mathcal{GM}$	$m_{\mathcal{GM},\mathcal{GW}}$	$m_{\mathcal{GM},\mathcal{GM}}$	$m_{\mathcal{GM},\mathcal{B}\cdot}$	$\mathcal{GM}$	$M_{\mathcal{GM},\mathcal{GW}}$	$M_{\mathcal{GM},\mathcal{GM}}$	$M_{\mathcal{GM},\mathcal{B}\cdot}$
$\mathcal{BM}$	$m_{\mathcal{BM},\mathcal{GW}}$	$m_{\mathcal{BM},\mathcal{GM}}$	$m_{\mathcal{BM},\mathcal{B}\cdot}$	$\mathcal{BM}$	$M_{\mathcal{BM},\mathcal{GW}}$	$M_{\mathcal{BM},\mathcal{GM}}$	$M_{\mathcal{BM},\mathcal{B}\cdot}$
$\mathcal{BW}$	$m_{\mathcal{BW},\mathcal{GW}}$	$m_{\mathcal{BW},\mathcal{GM}}$	$m_{\mathcal{BW},\mathcal{B}\cdot}$	$\mathcal{BW}$	$M_{\mathcal{BW},\mathcal{GW}}$	$M_{\mathcal{BW},\mathcal{GM}}$	$M_{\mathcal{BW},\mathcal{B}\cdot}$

sample mixing totals. For $i, i' > \tilde{I}$, we write

(8)
$$M_{ij,i'\cdot} = \frac{\left(\sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{ij}, \pi_{ij}; k\right) kN_{ij}\right) \left(\sum_{j'=1}^{J} \sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{i'j'}, \pi_{i'j'}; k\right) kN_{i'j'}\right)}{\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{ij}, \pi_{ij}; k\right) kN_{ij}} \alpha_{ij,i'\cdot},$$

where $\alpha_{ij,i'\cdot}$ gives the relative propensity of egos of type ij to have alters of type i' as opposed to alters of other types. Henceforth, we denote the vector of identifiable mixing parameters under this reparametrization by $\boldsymbol{\alpha}$.

Having reparametrized the model to ensure identifiability of unknown parameters, we can write the multinomial distribution of row ij of the matrix of sample mixing totals $\boldsymbol{m}_{ij} = \left(m_{ij,11}, \ldots, m_{ij,\tilde{I}J}, m_{ij,(\tilde{I}+1)\cdot}, \ldots, m_{ij,I\cdot}\right)$ given the sample degree distribution for egos of type ij $\boldsymbol{d}_{ij}$:

(9)
$$\boldsymbol{m}_{ij} | \boldsymbol{d}_{ij} \sim \mathcal{MN}\left(\sum_{k=0}^{K} kd_{ij,k}, \boldsymbol{p}_{\mathcal{M}}\left(\boldsymbol{\alpha}_{ij}, \boldsymbol{\rho}, \boldsymbol{\pi}; n_{ij}, \boldsymbol{N}\right)\right),$$

where

(10)
$$\boldsymbol{\alpha}_{ij} = \begin{cases} \left(\alpha_{ij,11}, \ldots, \alpha_{ij,IJ}\right) & \text{for } i \leq \tilde{I} \\ \left(\alpha_{ij,11}, \ldots, \alpha_{ij,\tilde{I}J}, \alpha_{ij,(\tilde{I}+1)\cdot}, \ldots, \alpha_{ij,I\cdot}\right) & \text{for } i > \tilde{I}. \end{cases}$$

The elements of the vector of multinomial probabilities

(11)
$$\boldsymbol{p}_{\mathcal{M}}\left(\boldsymbol{\alpha}_{ij}, \boldsymbol{\rho}, \boldsymbol{\pi}; n_{ij}, \boldsymbol{N}\right) = \big(\; p_{\mathcal{M},11}\left(\boldsymbol{\alpha}_{ij}, \boldsymbol{\rho}, \boldsymbol{\pi}; n_{ij}, \boldsymbol{N}\right), \ldots, p_{\mathcal{M},\tilde{I}J}\left(\boldsymbol{\alpha}_{ij}, \boldsymbol{\rho}, \boldsymbol{\pi}; n_{ij}, \boldsymbol{N}\right),$$
$$p_{\mathcal{M},(\tilde{I}+1)\cdot}\left(\boldsymbol{\alpha}_{ij}, \boldsymbol{\rho}, \boldsymbol{\pi}; n_{ij}, \boldsymbol{N}\right), \ldots, p_{\mathcal{M},I\cdot}\left(\boldsymbol{\alpha}_{ij}, \boldsymbol{\rho}, \boldsymbol{\pi}; n_{ij}, \boldsymbol{N}\right)\big)$$

are defined as

(12)

$$p_{\mathcal{M},i'j'}\left(\boldsymbol{\alpha}_{ij},\boldsymbol{\rho},\boldsymbol{\pi};n_{ij},\boldsymbol{N}\right) = \frac{\sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{i'j'},\pi_{i'j'};k\right)kN_{i'j'}}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{ij},\pi_{ij};k\right)kN_{ij}}\alpha_{ij,i'j'}$$

for $i, i' \leq \tilde{I}$ and for $i' > \tilde{I}$,

(13)

$$\boldsymbol{p}_{\mathcal{M},i'\cdot}\left(\boldsymbol{\alpha}_{ij},\boldsymbol{\rho},\boldsymbol{\pi};n_{ij},\boldsymbol{N}\right) = \begin{cases} \frac{\left(\sum_{j'=1}^{J}\alpha_{ij,i'j'}\sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{i'j'},\pi_{i'j'};k\right)kN_{i'j'}\right)}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=0}^{K} p_{\mathcal{D}}(\rho_{ij},\pi_{ij};k)kN_{ij}} & \text{for } i \leq \tilde{I} \\ \frac{\sum_{j'=1}^{J}\sum_{k=0}^{K} p_{\mathcal{D}}\left(\rho_{i'j'},\pi_{i'j'};k\right)kN_{i'j'}}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=0}^{K} p_{\mathcal{D}}(\rho_{ij},\pi_{ij};k)kN_{ij}}\alpha_{ij,i'\cdot} & \text{for } i > \tilde{I}. \end{cases}$$

Assuming independence across rows of the matrix of sample mixing totals, the model given by equations (6) and (9) yields the following likelihood

(14)

$$\mathcal{L}\left(\boldsymbol{\rho},\boldsymbol{\pi},\boldsymbol{\alpha};\boldsymbol{n},\boldsymbol{N}\right) = \left(\prod_{i\in 1,\ldots,I}\prod_{j\in 1,\ldots,J}\binom{n_{ij}}{d_{ij,0},\ldots,d_{ij,K}}\prod_{k=0}^{K}\left(p_{\mathcal{D}}\left(\rho_{ij},\pi_{ij};k\right)\right)^{d_{ij,k}}\right)$$

$$\left(\prod_{i\in 1,\ldots,I}\prod_{j\in 1,\ldots,J}\binom{\sum_{k=0}^{K}kd_{ij,k}}{m_{ij,\cdot 11},\ldots,m_{ij,\tilde{I}J},m_{ij,(\tilde{I}+1)\cdot},\ldots,m_{ij,I\cdot}}\right)\right)$$

$$\left(\prod_{i'=1}^{\tilde{I}}\prod_{j'=1}^{J} p_{\mathcal{M},i'j'}\left(\boldsymbol{\alpha}_{ij},\boldsymbol{\rho},\boldsymbol{\pi};n_{ij},\boldsymbol{N}\right)^{m_{ij,\cdot i'j'}}\right)$$

$$\left(\prod_{i'=\tilde{I}+1}^{I} p_{\mathcal{M},i'\cdot}\left(\boldsymbol{\alpha}_{ij},\boldsymbol{\rho},\boldsymbol{\pi};n_{ij},\boldsymbol{N}\right)^{m_{ij,i'\cdot}}\right),$$

where $p_{\mathcal{D}}\left(\rho_{ij},\pi_{ij};k\right)$ is as defined in equation (1) and elements of $\boldsymbol{p}_{\mathcal{M}}\left(\boldsymbol{\alpha}_{ij},\boldsymbol{\rho},\boldsymbol{\pi};n_{ij},\boldsymbol{N}\right)$ are defined in equations (12) and (13). Assuming independence across rows of the sample matrix of mixing totals is likely be valid when the sampled data represent a small proportion of egos in the network, as will often be the case when assessing RDS feasibility because the pilot data sample size will be small relative to size of the target population.

We maximize the likelihood from equation (14) subject to the following constraints, which follow from the assumption that edges are undirected and

the definitions of degree distributions and mixing totals:

$$\alpha_{ij,i'j'} = \alpha_{i'j',ij}, \tag{15}$$

$$\sum_{k=0}^{K} kD_{ij,k} = \sum_{i'=1}^{\tilde{I}} \sum_{j'=1}^{J} M_{ij,i'j'} + \sum_{i'=\tilde{I}+1}^{I} M_{ij,i'\cdot}, \tag{16}$$

$$\rho_{ij} > 0, \, 0 < \pi_{ij} < 1, \tag{17}$$

where the quantities $D_{ij,k}$, $M_{ij,i'j'}$ and $M_{ij,i'\cdot}$ are calculated from degree distribution parameters $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$, mixing parameters $\boldsymbol{\alpha}$ and known population sizes $\boldsymbol{N}$ using equations (2), (3) and (8). We obtain estimates of degree distribution parameters $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$ and mixing parameters $\boldsymbol{\alpha}$ using the Rsolnp package for constrained maximization in the R computing environment (Ye, 1987; R Core Team, 2013).

Model Extensions.

Stratification of Degree Distributions by Additional Nodal Characteristics. $h = 1, \ldots, H$ index characteristics by which degree distributions vary but mixing totals do not. For example, in our pilot data on LGB older adults, h indexes age group: younger, 50-64 years ($h = \mathcal{Y}$) or older, 65-80 years ($h = \mathcal{O}$). We stratify degree distributions by age because social network research suggests that degree is associated with age and age is observed for sampled egos (Cornwell et al., 2008; Fredriksen-Goldsen et al., 2015), however we do not stratify mixing totals by age because age is not observed for alters reported by sampled egos. We incorporate stratification of degree distributions by additional characteristics by applying the model for the population and sample degree counts D_{ij} and d_{ij} to D_{hij} and d_{hij} and by replacing degree count terms, $\sum_{k=0}^{k} p_{\mathcal{D}}(\rho_{ij}, \pi_{ij}; k) kN_{ij}$, with sums over the new index h, $\sum_{h=1}^{H} \sum_{k=0}^{k} p_{\mathcal{D}}(\rho_{hij}, \pi_{hij}; k) kN_{hij}$ in equations (14)-(17).

Partially Known Population Sizes. In the method described in the previous sections as in Admiraal and Handcock (2016), we assume that estimates of subpopulation sizes N_{hij} are available. However, this assumption is unlikely to hold in practice. In our pilot data on LGB older adults, estimates of the sizes of LGB older adult populations are available by gender identity and age group but not by sexual identity. To estimate of the sizes of LGB older adult populations by gender, age group and sexual identity, we again rely on the assumption of simple random sampling of egos and assume:

$$(n_{h1j}, \ldots, n_{hIj}) \sim \mathcal{MN}(n_{h\cdot j}, \boldsymbol{\theta}_{h\cdot j}), \tag{18}$$

where $\boldsymbol{\theta}_{h\cdot j} = (\theta_{h1j}, \ldots, \theta_{hIj})$ is such that θ_{hij} gives the proportion of egos of type hij among egos of type h and j in the population. We incorporate equation (18) into the likelihood given in equation (14), add constraints that $0 \leq \theta_{hlj} \leq 1$ to equation (17) and replace N_{hij} with $\theta_{hij}N_{h\cdot j}$ where N_{hij} appears in equation (14)-(17).

Non-Simple Random Sampling of Sample Egos. We assume that the sampled egos in the pilot data are a simple random sample throughout this paper because our pilot data on LGB older adults is a convenience sample for which the sampling frame is unknown. However, it is plausible that in many settings where researchers are interested in assessing RDS feasibility, the available pilot data may have been collected using a known sampling design for which sampling probabilities are known or can be estimated, e.g. the pilot data may be a previous RDS sample. In this case, the method for estimating the unknown network structure parameters $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$ and mixing parameters $\boldsymbol{\alpha}$ can be modified by generalizing our definition of $\boldsymbol{c}_{ijl}$:

$$\boldsymbol{c}_{ijl} = w_{ijl} \left(c_{ijl,11}, \ldots, c_{ijl,\tilde{I}J}, c_{ijl,(\tilde{I}+1)\cdot}, \ldots, c_{ijl,I\cdot} \right),$$

where $c_{ijl,i'j'}$ and $c_{ijl,i'\cdot} = \sum_{j'=1}^{J} c_{ijl,i'j'}$ are still the l-th sampled ego of type ij's reported number of alters of type $i'j'$ and i', respectively, and $l = 1, \ldots, n_{ij}$ and w_{ijl} is a sampling weight l-th sampled ego of type ij. Using this definition of $\boldsymbol{c}_{ijl}$, estimates of the unknown network structure parameters can be obtained using the method given earlier in this section.

3.2. *Simulation of Synthetic Populations.* Given estimates of the subpopulation sizes $\hat{\boldsymbol{N}}$, population degree distributions $\hat{\boldsymbol{D}}$ and mixing totals $\hat{\boldsymbol{M}}$, computed from estimates of degree distribution parameters $\hat{\boldsymbol{\rho}}$ and $\hat{\boldsymbol{\pi}}$, mixing parameters $\hat{\boldsymbol{\alpha}}$ and population proportions $\hat{\boldsymbol{\theta}}$ obtained by maximizing the likelihood given by modifications of equation (14) subject to the constraints given by equations (15)-(17) described in Section 3.1, we simulate synthetic networked populations.

We begin by initializing a random network with the degree distributions and subpopulation sizes given by $\hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{\pi}}$ and $\hat{\boldsymbol{\theta}}$ using functions in the igraph package for R (Gabor and Nepusz, 2006). Holding the degree of each ego constant, we obtain a random network with mixing totals consistent with $\hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{\pi}}$, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\theta}}$ using simulated annealing algorithms implemented in the ergm package (Handcock et al., 2008, 2013), a part of the statnet suite of packages for R (Handcock et al., 2003). Because many networks of size N with subpopulations $\hat{\boldsymbol{N}}$ have the same degree distributions and mixing totals, we use simulated annealing again to simulate 50 synthetic networked

populations with subpopulations $\hat{\boldsymbol{N}}$ holding the degree distributions and mixing totals constant at $\hat{\boldsymbol{D}}$ and $\hat{\boldsymbol{M}}$.

Although this approach is straightforward, difficulties arise in practice. First, this method scales poorly when the target population is large. Following Admiraal and Handcock (2016), we scale the networks down to the tractable size of $N = 5,000$ egos by dividing the subpopulation sizes by a constant that yields a total population size of $N = 5,000$ egos and recalculate the degree distributions and mixing totals using estimates $\hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{\pi}}$, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\theta}}$ and equations (2), (3) and (8). However, performance of RDS estimators can depend on the size of the population of interest. Specifically, RDS estimator biases can arise via finite population effects when an RDS sample is large relative to the corresponding population. We assess the presence of finite population effects incurred by artificially scaling down the networks to be more tractable by comparing RDS estimators that differ in their accounting of finite population bias, as suggested in Gile et al. (2015). Additionally, we repeat our RDS feasibility methodology on networks of size $4,000$ and $6,000$ as a sensitivity analysis. If RDS estimators that differ in their accounting of finite population bias perform similarly and if RDS estimator performance varies little as network size varies from $4,000$ to $6,000$, we assume that the performance of RDS estimators for the scaled population size of $N = 5,000$ egos is indicative of the performance of RDS estimators for the larger, true population size. We expect that networks of size $5,000$ should suffice in many settings, as recent research on finite population effects suggests that they are negligible when the RDS sample makes up less than 20% of the total population. However, if evidence of finite population effects is present, we recommend using larger synthetic networks, despite the increased computational burden.

A second difficulty is that estimated degree counts are not integers and therefore are not realizable. In practice, we round the estimated degree counts to the nearest integer. However when degree counts are small, rounding may result in degree distributions that cannot be initialized and network initialization may result in a network with a degree distribution that differs from the estimated degree distribution. To address this issue, we take the network obtained using functions in the `igraph` package with a degree distribution that is closest to the estimated degree distribution. We also address this problem by enforcing a degree distribution upper bound K. We choose this upper bound by recognizing a trade-off between stability of estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ and degree distribution realizability. Small values of K yield few small degree counts, however they also yield estimates of mixing parameters, $\boldsymbol{\alpha}$, and population proportion parameters, $\boldsymbol{\theta}$, that are sensitive to changes

in K. In contrast, large values of K yield many small degree counts that can lead to synthetic networked populations with degree distributions far from the estimated degree distributions, but stable estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$. Therefore, we recommend choosing the largest value of K such that the degree distribution of the synthetic networked populations is close to the estimated degree distribution while estimates $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ have stabilized.

Before moving on to modeling the characteristic(s) of interest, we note that it is possible to simulate synthetic networks of size N with stochastic subpopulations $\hat{\boldsymbol{N}}$, degree distributions $\hat{\boldsymbol{D}}$ and mixing totals $\hat{\boldsymbol{M}}$ that reflect the uncertainty of estimates $\hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{\pi}}$ and $\hat{\boldsymbol{\theta}}$. We could draw values of $\boldsymbol{\rho}$, $\boldsymbol{\pi}$ and $\boldsymbol{\theta}$ from a normal distribution centered at $\hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{\pi}}$ and $\hat{\boldsymbol{\theta}}$ with variance-covariance matrix given by the inverse of the Fisher information matrix corresponding to Equations (14) and (15). For each value of $\boldsymbol{\rho}$, $\boldsymbol{\pi}$ and $\boldsymbol{\theta}$, synthetic populations could be simulated according to the procedure described at the beginning of this section. Population sizes were too large for this to be computationally tractable for the data considered in this paper, however for smaller population sizes this is a natural extension to the synthetic network simulation procedure.

Modeling the Characteristic(s) of Interest. Recall that our goal is to simulate RDS for estimating certain characteristics of the target population. When these characteristics are not indexed by i, j or h, we need to assign them to egos in the synthetic networked populations. To simplify interpretation of the results of the feasibility assessment, we fix the number of egos with each characteristic not indexed by i, j or h across synthetic networked populations. We then randomly assign these characteristics to egos with probabilities estimated from logistic regression models of the characteristics of interest fit to the pilot data.

3.3. *Simulation of RDS.* On each of the 50 synthetic networked populations, we simulate RDS 500 times. Each RDS sample begins with 10 egos selected as seeds at random using a sampling strategy that reflects the features of the target population and plans for seed recruitment. We note the common strategy of selecting seeds with probability proportional to degree is is favorable to the Volz-Heckathorn estimator and should only be used if it is a reasonable assumption given the target population and plans for seed recruitment (Gile and Handcock, 2010). Otherwise simulations may suggest unrealistically good performance of the Volz-Heckathorn estimator. Egos of type hij drawn for inclusion in the study "redeem coupons" with probability $\hat{\pi}_{hij}^{\mathcal{C}}$ and "distribute coupons" with probability $\hat{\pi}_{hij}^{\mathcal{R}}$. These probabilities

are estimated from the pilot data by fitting logistic regressions that model measures of willingness to redeem and distribute coupons as a function of egos' attributes, e.g. characteristics indexed by h, i and j or degree. The process iterates without replacement until the desired RDS sample size is obtained, where the desired RDS sample size is the RDS sample size the researcher is considering using in practice and wants to assess the feasibility of. If the RDS sample "dies out" before achieving the desired RDS sample size, additional seeds are added and RDS is continued until the desired RDS sample size is obtained, as is often done in practice.

As described in the Introduction, we consider the sample mean, the Salganik-Heckathorn (SH) and the Volz-Heckathorn (VH) estimators as point estimators and compute 95% confidence intervals using exact binomial confidence intervals for the sample mean, the bootstrap procedure introduced in Salganik (2006) for the SH and VH estimators (Salganik and Heckathorn, 2004; Volz and Heckathorn, 2008; Gile, 2011). For a thorough review of the point estimators and explicit expressions for each, see Tomas and Gile (2011).

3.4. *RDS Feasibility Assessment.* As described in Section 1.2, we assess RDS feasibility by comparing thresholds for effective sample sizes, confidence interval widths and confidence interval coverage rates τ_S, τ_W and τ_C to the effective sample sizes, confidence interval widths and confidence interval coverage rates of RDS estimators across simulations.

4. Application to CAP Data. We apply the methodology described in Section 3 to pilot data on LGB older adults described in Section 2 to assess the feasibility of using RDS to estimate the proportion male, proportion bisexual, proportion depressed and proportion infected with HIV/AIDS in three spatially distinct populations. First, we present estimates of subpopulation proportions $\boldsymbol{\theta}$, degree distributions $\boldsymbol{\pi}^{\mathcal{D}}$ derived from degree distribution parameters $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$, and mixing total parameters $\boldsymbol{\alpha}$. We describe how these estimates inform our our expectations of RDS feasibility for estimating characteristics of these populations. Second, we present the performance of RDS estimators for estimating the proportion male, proportion bisexual, proportion depressed and proportion infected with HIV/AIDS from a simulation study. We conclude by assessing feasibility for estimating these characteristics in these populations. To assess feasibility, we set the following thresholds for effective sample size for an RDS sample of 300 subjects and coverage, according to the general standards we describe in the Introduction: $\tau_S = 100$, $\tau_W = \inf$ and $\tau_C = 80\%$.

4.1. *Estimation of Network Population Parameters.* Existing research on LGB older populations and exploratory analyses suggest that degree varies by gender identity and age (Erosheva et al., 2016). Consequently, we stratify degree distributions by gender identity and age group. Existing evidence for variation in network diversity by sexual and gender identity leads us to stratify mixing totals by gender and sexual identity (Erosheva et al., 2016). We choose a maximum degree K of 45 based on the guidelines given in Section 3.2. Evidence supporting this choice is given in Section 2 of the supplement.
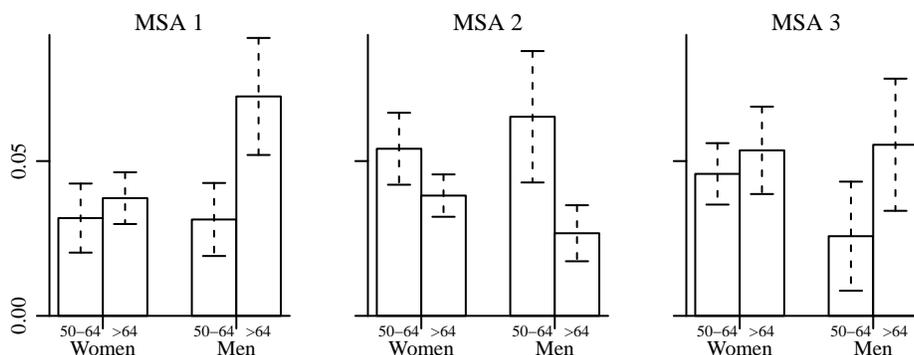

Fig 2: Estimates of the proportion bisexual by gender and age group $\boldsymbol{\theta}$ for each MSA. Dashed bars represent standard errors.

Figure 2 shows estimates and standard errors for subpopulation proportions $\boldsymbol{\theta}$. Figure 3 shows estimates of the degree distributions $\boldsymbol{\pi}^{\mathcal{D}}$ derived from estimates of the degree distributions parameters $\boldsymbol{\rho}$ and $\boldsymbol{\pi}$ by gender identity and age group within each MSA. Figure 4 shows estimates of mixing parameters $\boldsymbol{\alpha}$ which measure preferential mixing, for the three MSAs.

Variation in degree distribution by gender identity and age group indicates differential activity with respect to gender identity and age group, while preferential mixing with respect to gender identity indicates homophily with respect to gender identity. Differential activity with respect to gender identity is defined as the ratio of the average degree of male egos to the average degree of female egos and homophily with respect to gender identity is defined as the ratio of observed edges between egos of the same gender identity to the expected number of edges between egos and alters of the same gender identity under random mixing (Gile and Handcock, 2015). Homophily and differential activity with respect to gender identity and age group can affect the performance of RDS estimators of characteristics asso-
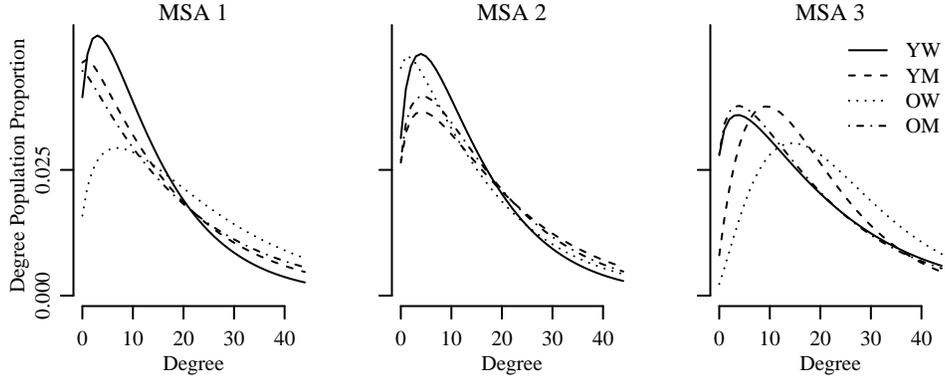
Fig 3: Estimates of degree distributions by gender and age group constructed from estimates of negative binomial parameters ρ and π. Different line types correspond to different age groups and gender identities.

ciated with gender identity and age group (Gile and Handcock, 2010; Tomas and Gile, 2011). This suggests that RDS *may* not be feasible for estimating the proportion male, the proportion under 65 and characteristics associated with gender identity or age group in these populations.

4.2. *RDS Feasibility Assessment.* We assess RDS feasibility for estimating the proportion male, proportion bisexual, proportion depressed and proportion infected with HIV/AIDS in MSAs 1, 2 and 3 by simulating RDS on synthetic networked populations consistent with the estimated network structures described in Section 4.1. For each MSA, we simulate 50 synthetic networked populations of $N = 5,000$ egos with subpopulation sizes, degree distributions and mixing totals given by estimated population proportions $\hat{\boldsymbol{\theta}}$, degree distribution parameters $\hat{\boldsymbol{\pi}}^{\mathcal{D}}$ and mixing parameters $\hat{\boldsymbol{\alpha}}$.

As described in Section 3.2, we fix the number of depressed and HIV/AIDS infected egos across synthetic networked populations corresponding to the same MSA. We then assign depression and HIV/AIDS status to fixed numbers of individual nodes at random with probabilities estimated from logistic regression models trained on the CAP data within each MSA. The logistic regression models have sexual identity, gender identity, age group, an indicator for zero network size (0 alters), the total number of alters and the number of alters of the same gender and sexual identity as covariates. Regression coefficient estimates are shown in Section 3 of the supplement. Table 3 shows the proportion male, proportion bisexual, proportion depressed and proportion infected with HIV/AIDS in the synthetic networked populations
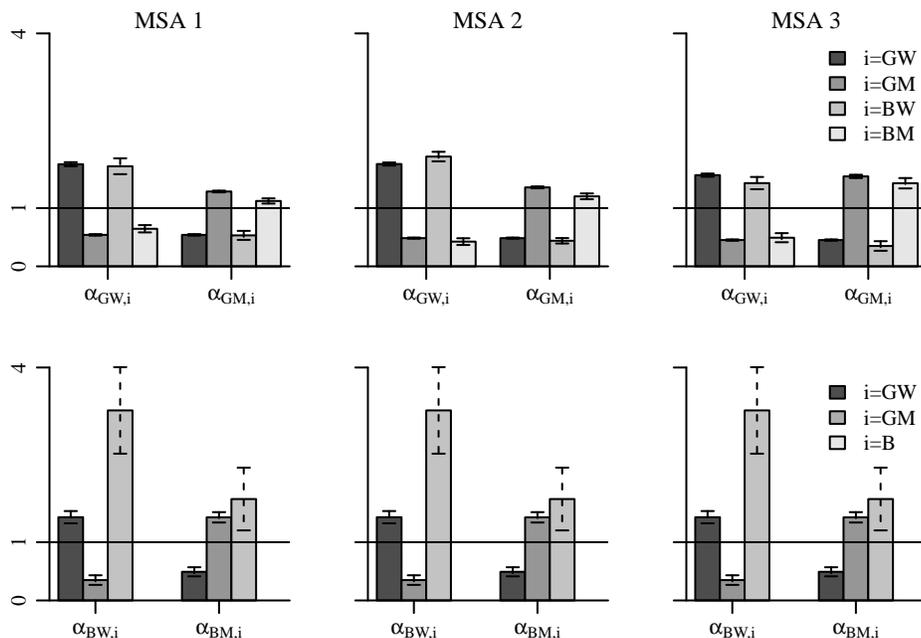
Fig 4: Estimates of mixing parameters $\boldsymbol{\alpha}$. For each estimated mixing parameter $\hat{\alpha}_{ij}$, i is indexed on the horizontal axis and j is indexed by color. For instance, the estimate $\hat{\alpha}_{\mathcal{GW},\mathcal{GW}}$, is given by the first bar from the left in panel (a). $\alpha_{ij} = 1$ reflects no preferential mixing, i.e. egos of type i are more as likely to have alters of type j as alters of other types. $|\alpha_{ij}| > 1$ indicates preferential mixing, i.e. egos of type i are more or less likely to have alters of type j than other types. Dashed bars represent standard errors.

representing each MSA.

Lastly, we assign probabilities representing willingness to participate and willingness to refer others $\boldsymbol{\pi}^{\mathcal{C}}_{hij}$ and $\boldsymbol{\pi}^{\mathcal{R}}_{hij}$ to egos in the synthetic networked populations using logistic regression models trained on the CAP data within each MSA. These logistic regression models also have sexual identity, gender identity, age group, an indicator for zero network size (0 alters), the total number of alters and the number of alters of the same gender and sexual identity as covariates. Regression coefficient estimates are shown in Section 4 of the supplement.

We simulate RDS on these synthetic networked populations as described in Section 3.3. We sample seeds with probability proportional to degree. We believe that this is a reasonable assumption in this context where seeds are

	% Male	% Bisexual	% Depressed	% Infected with HIV/AIDS
MSA 1	61.50	4.00	43.36	20.78
MSA 2	54.66	4.48	23.00	6.26
MSA 3	52.50	4.86	24.40	13.06

TABLE 3

Characteristics of Interest in Synthetic Networks

recruited from agencies serving LGB older adults. LGB older adults tend to report high rates of social isolation, but those associated with such agencies are likely to have higher-than-average degree. To assess RDS feasibility, we simulate RDS 500 times for a desired RDS sample size of 300, a common RDS sample size chosen in practice. Before assessing RDS feasibility, we use the procedure described in Section 3.2 to confirm the absence of finite population bias. Evidence for the absence of population bias is shown in Section 5 of the supplement.

Estimating the Proportion Male. Table 4 shows the effective sample sizes, confidence interval widths and confidence interval coverage rates for estimating the proportion male. With the exception of the sample mean in MSA 1, the effective sample size for estimators of the proportion male is less than 100. We also observe that only the VH estimators provide coverage over 80%. These results indicate that RDS is not feasible for estimating the proportion male in any of the three MSAs.

	Sample Mean			SH			VH		
	ESS	WI	CVR	ESS	WI	CVR	ESS	WI	CVR
MSA 1	**105.76**	**0.11**	0.71	65.17	**0.13**	0.74	64.97	**0.22**	**0.92**
MSA 2	87.95	**0.11**	0.41	58.85	**0.14**	0.75	57.87	**0.23**	**0.90**
MSA 3	74.15	**0.11**	0.61	61.57	**0.13**	0.70	55.86	**0.23**	**0.91**

TABLE 4

Effective sample sizes (ESS), 95% confidence interval widths (WI) and coverage (CVR) for estimating the proportion male. Bold entries satisfy our feasibility criteria.

Estimating the Proportion Bisexual. Table 5 shows the effective sample sizes, confidence interval widths and confidence interval coverage rates for estimating the proportion bisexual. The effective sample size exceeds 100 for the sample mean, VH estimators in all three MSAs. We also observe that the same three estimators provide coverage over 80%. These results indicate that RDS may be feasible for estimating the proportion bisexual in any of the three MSAs, providing that the sample mean or VH estimator is used.

Estimating the Proportion Depressed. Table 6 shows the effective sample sizes, confidence interval widths and confidence interval coverage rates for estimating the proportion depressed. As with the proportion bisexual, the

	Sample Mean			SH			VH		
	ESS	*WI*	*CVR*	*ESS*	*WI*	*CVR*	*ESS*	*WI*	*CVR*
MSA 1	**311.71**	**0.04**	**0.92**	56.70	**0.06**	0.76	**111.10**	**0.07**	**0.83**
MSA 2	**303.39**	**0.04**	**0.93**	90.61	**0.06**	0.79	**114.16**	**0.08**	**0.85**
MSA 3	**274.83**	**0.05**	**0.93**	**110.15**	**0.06**	0.78	**124.05**	**0.08**	**0.86**

TABLE 5

Effective sample sizes (ESS), 95% confidence interval widths (WI) and coverage (CVR) for estimating the proportion bisexual. Bold entries satisfy our feasibility criteria.

effective sample size exceeds 100 for all four estimators and all three MSAs. We also observe that coverage is almost always over 80%, with the exception of the coverage of the sample mean in MSAs 2 and 3. These results indicate that RDS may be feasible for estimating the proportion depressed in any of the three MSAs, providing that the SH or VH estimator are used in MSAs 2 and 3.

	Sample Mean			SH			VH		
	ESS	*WI*	*CVR*	*ESS*	*WI*	*CVR*	*ESS*	*WI*	*CVR*
MSA 1	**299.52**	**0.11**	**0.94**	**107.58**	**0.13**	**0.83**	**109.28**	**0.23**	**0.98**
MSA 2	**324.61**	**0.09**	0.77	**110.68**	**0.12**	**0.84**	**112.80**	**0.19**	**0.95**
MSA 3	**344.81**	**0.09**	0.48	**111.69**	**0.14**	**0.90**	**113.64**	**0.21**	**0.97**

TABLE 6

Effective sample sizes (ESS), 95% confidence interval widths (WI) and coverage (CVR) for estimating the proportion depressed. Bold entries satisfy our feasibility criteria.

Estimating the Proportion Infected with HIV/AIDS. Table 7 the effective sample sizes, confidence interval widths and confidence interval coverage rates for estimating the proportion infected with HIV/AIDS. As with the proportion bisexual and depressed, the effective sample size exceeds 100 in most cases, with the exception of the SH estimator in MSAs 1 and 2. We also observe that coverage is almost always over 80%, with the exception of the coverage of the sample mean in MSAs 3 and the SH estimator in MSA 1. These results indicate that RDS may be feasible for estimating the proportion infected with HIV/AIDS in any of the three MSAs, providing that the appropriate estimators are used.

	Sample Mean			SH			VH		
	ESS	*WI*	*CVR*	*ESS*	*WI*	*CVR*	*ESS*	*WI*	*CVR*
MSA 1	**218.05**	**0.09**	**0.88**	99.85	**0.10**	0.79	**101.20**	**0.42**	**0.99**
MSA 2	**304.95**	**0.05**	**0.85**	97.05	**0.08**	**0.82**	**105.56**	**0.31**	**0.99**
MSA 3	**242.32**	**0.07**	0.74	**101.04**	**0.11**	**0.87**	**103.60**	**0.30**	**1.00**

TABLE 7

Effective sample sizes (ESS), 95% confidence interval widths (WI) and coverage (CVR) for estimating the proportion infected with HIV/AIDS. Bold entries satisfy our feasibility criteria.

That being said, we do observe substantial over-coverage of the VH estimators, exceeding 95%. If we examine confidence interval widths for these estimators, we observe that the the widths for the VH 95% confidence intervals are very large relative to the 95% confidence intervals given by the other estimators. Given that we do not set standards for improvement for estimating these characteristics of these populations because there are so few existing estimates, this does not affect our assessment of feasibility. However, it does raise the philosophical question of whether or not a very wide 95% confidence interval can still be useful when no alternative estimates from previous studies are available. In this case, confidence intervals of width 0.4 as observed in MSA 1, may be practically useless.

Interpreting RDS Estimator Performance. We explain the relative performance of the RDS estimators shown in Tables 4-7 by examining homophily in the synthetic networked populations and recruitment effectiveness of the simulated RDS process. Recruitment effectiveness is the ratio of the average number of alters recruited by egos who do not possess the characteristic of interest relative to the average number of alters recruited by egos who do possess the characteristic of interest (Gile et al., 2015); it is a composite measure of differential activity and differential coupon distribution and redemption behavior. Although homophily and differential recruitment are not separately identified from standard RDS data alone in practice (Crawford et al., 2017), we are able to discuss both separately here because homophily is explicitly measured in the pilot data on LGB older adults by asking sampled egos about the characteristics of their alters. As noted in Section 4.1, homophily and differential activity with respect to the characteristic of interest can affect the performance of RDS estimators (Gile and Handcock, 2010; Tomas and Gile, 2011). High homophily in particular tends to produce more variable RDS estimators over repeated RDS samples. Table 8 shows homophily and differential recruitment effectiveness with respect to all four characteristics of interest in all three MSAs.

	Homophily				Diff. Recruitment Eff.			
	Gen.	Bi.	Dep.	HIV	Gen.	Bi.	Dep.	HIV
MSA 1	1.83	1.05	1.00	1.08	0.92	1.10	0.99	0.96
MSA 2	2.08	1.04	1.00	1.02	0.92	1.15	1.03	0.98
MSA 3	2.20	1.08	1.00	1.07	1.06	0.97	1.03	1.05

TABLE 8

Homophily and Differential Recruitment Effectiveness by Gender, Bisexuality, Depression and HIV/AIDS

We observe strong evidence for homophily with respect to gender identity

and little evidence for homophily with respect to sexual identity, depression status and HIV/AIDS infection status, which explains the infeasibility of RDS for estimating the proportion male. This suggests that RDS estimator performance of the proportion male may not improve if a larger RDS sample is collected. Additionally, this suggests that gay and bisexual men and lesbian and bisexual women have somewhat separate social spheres that may be best served by separate RDS studies combined using estimates of the proportion male and female obtained from a much more costly large scale probability sample, e.g. the Gallup Daily Tracking Survey. We observe little evidence for differential recruitment effectiveness and no evidence of systematic differences in RDS feasibility with respect to differential recruitment effectiveness.

5. Summary and Discussion. In this paper, we develop a simulation-based framework for assessing the feasibility of using RDS to estimate specific characteristics of target populations given pilot data, in the form of aggregated egocentric data and estimates of subpopulation sizes within the target population. Such a framework is currently unavailable, leaving researchers to plan costly studies based on qualitative feasibility assessment criteria that may fail to detect network structures and coupon distribution behaviors that yield poor RDS estimator performance (Johnston et al., 2010; Kogan et al., 2011). To explicitly characterize feasibility, we consider the effective sample size of point estimates and the width and coverage of 95% confidence intervals and explain how these quantities can be interpreted by a researcher considering conducting an RDS study. In introducing a framework for assessing RDS feasibility, we extend Admiraal and Handcock (2016)'s work on simulating social networks consistent with aggregated, egocentric sample data by using a computationally simple and sufficiently flexible negative binomial distribution model for broad degree distributions, specifying rectangular matrices of sample mixing totals and allowing for unobserved subpopulation sizes. Rectangular specification for the sample mixing matrices allows us to accommodate the common scenario when more information is available on egos than on alters.

With respect to LGB older populations, we find that using RDS with 10 seeds, 2 coupons and a desired sample size of 300 subjects is infeasible for estimating the proportion male in all three target populations and feasible for estimating the proportion bisexual, proportion depressed and proportion infected with HIV/AIDS. We also find that effective sample sizes, interval widths and coverage for estimators estimating the same outcome can vary by MSA, e.g. effective sample sizes and coverage of 95% confidence inter-

vals are systematically higher for all four RDS estimators of the proportion male in MSA 1 compared to MSAs 2 and 3. These results not only illustrate the possible feasibility of RDS in these LGB older populations for estimating the proportion bisexual, proportion depressed and proportion infected with HIV/AIDS, but also demonstrate that RDS feasibility assessment is context-specific. Feasibility of RDS in one MSA for one characteristic of interest provides limited information about the feasibility of RDS in other MSAs or for other characteristics of interest. Specifically, homophily with respect to gender yields poor estimation of the proportion male in all three populations but does not affect estimation of the proportion bisexual, depressed or infected with HIV/AIDS.

We caution that acceptable performance of RDS estimators in simulations does not guarantee acceptable performance of RDS estimators in practice. In many ways, our method for estimating network parameters and simulating RDS represents an upper bound on the performance of RDS. Although it is realistic in that it incorporates what is known from the pilot data on social network structures and participation and recruitment tendencies, it is unrealistic in many ways. For example, it assumes the pilot data is a simple random sample from the target population, that RDS sample seeds are selected with probability proportional to degree and that reported coupon distribution and redemption behaviors are consistent with observed coupon distribution and redemption behaviors. As a result, we recommend interpreting poor performance of RDS estimators in simulations as a red flag, while we do not recommend interpreting acceptable performance of RDS estimators in simulations as a guarantee that RDS will necessarily perform well in practice. We therefore recommend that researchers interested in assessing RDS feasibility collect pilot data, then implement the methods introduced in this paper. If feasibility is not demonstrated, we recommend against proceeding with RDS and in favor of pursuing a more expensive large scale probability sample if possible or alternative non-probability sampling methods like time-location sampling if not (Gile and Handcock, 2010). If RDS is found to be feasible, as with any statistical approximation, it is important for the researcher to reflect on any critical concerns that may exist beyond the scope of the available data and methods.

Although beyond the scope of the paper, our conceptual framework also allows interested researchers to explore feasibility of RDS with respect to variations in RDS designs such as changing the number and characteristics of seeds, the number of coupons, participation rates, and the desired RDS sample size. Future work could consider extending the RDS feasibility framework presented in this paper to evaluating multiple RDS designs by relying

on ideas analogous to power analysis (Cohen, 1988) or diagnostic approaches developed in the context of assessment and ranking of probabilistic weather forecasts (Gneiting et al., 2007).

To ensure that the tools for implementing our simulation-based framework are available to practitioners, we made R code and a synthetic data dataset available as supplementary material. The synthetic dataset was constructed to resemble one of the three LGB older populations studied in this paper. The R code is accompanied by a step-by-step guide which walks through estimating the social network structure, simulating synthetic networked populations and simulating RDS. The guide includes comments on where modifications could be made to accommodate alternative seed selection mechanisms.

SUPPLEMENTARY MATERIAL

Supplemental tables, figures, details and discussion. (DOI: ????; .zip). The file contains a document titled "Supplement to 'A simulation-based framework for assessing the feasibility of respondent-driven sampling for estimating characteristics in populations of lesbian, gay and bisexual older adults,"' which includes the five sections referenced in this paper. The file also contains the source code for a package for R that includes code written by the authors and synthetic data. A vignette titled "Synthetic data example of methods used in 'A simulation-based framework for assessing the feasibility of respondent-driven sampling for estimating characteristics in populations of lesbian, gay and bisexual older adults"' implements the methods used in this paper for the synthetic data. Additionally, a read me file is included with instructions for installing the R package and accessing the vignette.

References.

Admiraal, R. and M. S. Handcock (2016). Modeling Concurrency and Selective Mixing in Heterosexual Partnership Networks With Applications to Sexually Transmitted Diseases. *Annals of Applied Statistics 10*(4), 2021–2046.

Andresen, E. M., J. A. Malmgren, W. B. Carter, and D. L. Patrick (1994). Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *American Journal of Preventive Medicine 10*(2), 77–84.

Barash, V. D., C. J. Cameron, M. W. Spiller, and D. D. Heckathorn (2016). Respondent-Driven Sampling Testing Assumptions: Sampling with Replacement. *Journal of Official Statistics 32*(1), 29–73.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Routledge.

Cornwell, B., E. O. Laumann, and L. P. Schumm (2008). The Social Connectedness of Older Adults: A National Profile. *American Sociological Review 73*(2), 185–203.

Crawford, F. W., P. M. Aronow, L. Zeng, and J. Li (2017). Identification of Homophily and Preferential Recruitment in Respondent-Driven Sampling. *American Journal of Epidemiology (in press)*.

Erosheva, E. A., H.-J. Kim, C. Emlet, and K. I. Fredriksen-Goldsen (2016). Social Networks of Lesbian, Gay, Bisexual, and Transgender Older Adults. *Research on Aging 38*(1), 98–123.

Fredriksen-Goldsen, K. I., C. A. Emlet, H.-J. Kim, A. Muraco, E. A. Erosheva, J. Goldsen, and C. P. Hoy-Ellis (2013). The Physical and Mental Health of Lesbian, Gay Male, and Bisexual (LGB) Older Adults: The Role of Key Health Indicators and Risk and Protective Factors. *The Gerontologist 53*(4), 664–675.

Fredriksen-Goldsen, K. I., H.-J. Kim, C. Shiu, J. Goldsen, and C. A. Emlet (2015). Successful Aging Among LGBT Older Adults: Physical and Mental Health-Related Quality of Life by Age Group. *The Gerontologist 55*(1), 154–168.

Fredriksen-Goldsen, K. I. and A. Muraco (2010). Aging and Sexual Orientation: A 25-Year Review of the Literature. *Research on Aging 32*(3), 372–413.

Gabor, C. and T. Nepusz (2006). The igraph software package for complex network research. *InterJournal, Complex Systems 1695*(5), 1–9.

Gates, G. J. (2013). LGBT Parenting in the United States. Technical report, UCLA: The Williams Institute.

Gile, K. J. (2011). Improved Inference for Respondent-Driven Sampling Data With Application to HIV Prevalence Estimation. *Journal of the American Statistical Association 106*(493), 135–146.

Gile, K. J. and M. S. Handcock (2010). Respondent-Driven Sampling: An Assessment of Current Methodology. *Sociological Methodology 40*(1), 285–327.

Gile, K. J. and M. S. Handcock (2015). Network model-assisted inference from respondent-driven sampling data. *Journal of the Royal Statistical Society. Series A: Statistics in Society 178*(3), 619–639.

Gile, K. J., L. G. Johnston, and M. J. Salganik (2015). Diagnostics for respondent-driven sampling. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 178*(1), 241–269.

Gneiting, T., F. Balabdaoui, and A. E. Raftery (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society. Series B: Statistical Methodology 69*(2), 243–268.

Goel, S. and M. J. Salganik (2010). Assessing respondent-driven sampling. *Proceedings of the National Academy of Sciences 107*(15), 6743–7.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, and M. Morris (2003). statnet: A suite of R packages for the statistical modeling of social networks. Software library.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, and M. Morris (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software 24*(3), 1–29.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, and M. Morris (2013). *ergm: Fit, simulate and analyze exponential-family models for networks*. Statnet Project.

Heckathorn, D. D. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. *Social Problems 44*(2), 174–199.

Johnston, L. G., S. Whitehead, M. Simic-Lawson, and C. Kendall (2010). Formative research to optimize respondent-driven sampling surveys among hard-to-reach populations in HIV behavioral and biological surveillance: lessons learned from four case

studies. *AIDS Care 22*(6), 784–792.

Kogan, S. M., C. Wejnert, Y.-F. Chen, G. H. Brody, and L. M. Slater (2011). Respondent-Driven Sampling With Hard-to-Reach Emerging Adults: An Introduction and Case Study With Rural African Americans. *Journal of Adolescent Research 26*(1), 30–60.

Li, X. and K. Rohe (2015). Central limit theorems for network driven sampling. *ArXiv preprint. arXiv:1509.04704*.

Lohr, S. L. (2010). *Sampling: Design and Analysis* (Second ed.). Boston, MA: Brooks/Cole.

Lu, X., L. Bengtsson, T. Britton, M. Camitz, B. J. Kim, A. Thorson, and F. Liljeros (2012). The sensitivity of respondent-driven sampling. *Journal of the Royal Statistical Society. Series A: Statistics in Society 175*(1), 191–216.

Malekinejad, M., L. G. Johnston, C. Kendall, L. R. F. S. Kerr, M. R. Rifkin, and G. W. Rutherford (2008). Using Respondent-Driven Sampling Methodology for HIV Biological and Behavioral Surveillance in International Settings: A Systematic Review. *AIDS and Behavior 12*(S), 105–130.

McCreesh, N., S. D. W. Frost, J. Seeley, J. Katongole, M. N. Tarsh, R. Ndunguse, F. Jichi, N. L. Lunel, D. Maher, L. G. Johnston, P. Sonnenberg, A. J. Copas, R. J. Hayes, and R. G. White (2012). Evaluation of respondent-driven sampling. *Epidemiology 23*(1), 138–147.

Merli, M. G., J. Moody, J. Smith, J. Li, S. Weir, and X. Chen (2015). Challenges to recruiting population representative samples of female sex workers in China using Respondent Driven Sampling. *Social Science and Medicine 125*, 79–93.

R Core Team (2013). R: A Language and Environment for Statistical Computing.

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement 1*(3), 385–401.

Rohe, K. (2015). Network driven sampling; a critical threshold for design effects. *ArXiv preprint. arXiv:1505.05461*.

Salganik, M. J. (2006). Variance Estimation, Design Effects, and Sample Size Calculations for Respondent-Driven Sampling. *Journal of Urban Health 83*(7), 98–112.

Salganik, M. J. and D. D. Heckathorn (2004). Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling. *Sociological Methodology 34*, 193–239.

Tomas, A. and K. J. Gile (2011). The Effect of Differential Recruitment, Non-response and Non-recruitment on Estimators for Respondent-Driven Sampling. *Electronic Journal of Statistics 5*, 899–934.

Verdery, A. M., T. Mouw, S. Bauldry, and P. J. Mucha (2015). Network structure and biased variance estimation in respondent driven sampling. *PLoS ONE 10*(12), 1–27.

Volz, E. and D. D. Heckathorn (2008). Probability Based Estimation Theory for Respondent Driven Sampling. *Journal of Official Statistics 24*(1), 79–97.

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

Ye, Y. (1987). *Interior algorithms for linear, quadratic, and linearly constrained non-linear programming.* Ph.d., Stanford University.

Zea, M. C. (2010). Reaction to the Special Issue on Centralizing the Experiences of LGB People of Color in Counseling Psychology. *The Counseling Psychologist 38*(3), 425–433.