SUPPLEMENT TO "INFERENCE FOR SOCIAL NETWORK MODELS FROM EGOCENTRICALLY-SAMPLED DATA, WITH APPLICATION TO UNDERSTANDING PERSISTENT RACIAL DISPARITIES IN HIV PREVALENCE IN THE US"

By Pavel N. Krivitsky and Martina Morris

Throughout this document, references to sections and equations beginning with a number (as opposed to a letter) refer to the main article (Krivitsky and Morris, 2016).

APPENDIX A: COMPUTATIONALLY EFFICIENT APPROXIMATION USING NETWORK SIZE ADJUSTMENT

Here, we give details for the computational approximation mentioned in Section 5.2 and derive its properties.

Krivitsky, Handcock, and Morris (2011) suggested an approach for a network-size-invariant parametrization for some ERGMs for undirected graphs, where a network of size $|N|$ is modeled with an offset term, i.e.,

$$\Pr_{\boldsymbol{g}}(\boldsymbol{Y}=\boldsymbol{y};\boldsymbol{x}_N,\boldsymbol{\theta}) \equiv \frac{\exp\{-\log(|N|)|\boldsymbol{y}|+\boldsymbol{\theta}^\top \boldsymbol{g}(\boldsymbol{y},\boldsymbol{x}_N)\}}{\kappa_{\boldsymbol{g}}(\boldsymbol{\theta},\boldsymbol{x}_N)}, \boldsymbol{y}\in 2^{\mathbb{Y}(N)}. \tag{A.1}$$

This adjustment works, particularly, for network processes that fulfill certain heuristics: locality, in that as the network size changes, an individual actor's egocentric view of the network does not, on average, change; and stable degree distribution and per-capita mixing, in that the distribution of the number of ties an actor has (and the distribution of attributes of those to whom the actor has ties) remain stable as network size changes, provided the composition is preserved. For network processes and ERGM terms fulfilling this, networks having similar structure and composition but different sizes produced the same parameter estimates after the network size adjustment. They demonstrated this rigorously for dyadic-independent ERGM terms and by simulation for degree distribution terms. (Hunter et al., 2008)

This finding suggests a straightforward computational shortcut: instead of constructing the full population network over actors N, one can construct a "scaled-down" version $N' \subseteq N$ having the same composition (distribution of $\boldsymbol{x}$) and large enough for the estimates to have asymptoted. Fitting (A.1)

with N replaced by N' and $\tilde{\boldsymbol{g}}(e_S)$ replaced by $\tilde{\boldsymbol{g}}(e_S) \times |N'|/|N|$ would then yield $\tilde{\boldsymbol{\theta}}^{N'} \approx \tilde{\boldsymbol{\theta}}^{N}$.

A.1. Requirements for the adjustment. By design, the network composition is fixed, with only size changing, so for the adjustment to work in our case, the ERGM must be *local* (which, in this case, holds by construction as described in Section 3.1) and degree distribution and per-capita mixing must be stable. In the context of egocentric ERGMs, this can be operationalized as the distribution of individual measurements $\boldsymbol{h}(\boldsymbol{e}_i)$ being unaffected by the network size. This is a property of the model, not of the data, and from the perspective of the model, this requires that,

$$\lim_{|N'|\to\infty} |N'|^{-1} \boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\boldsymbol{\theta}) = \boldsymbol{\lambda}_{\boldsymbol{g}}(\boldsymbol{\theta}), \tag{A.2}$$

with $\boldsymbol{\lambda}_{\boldsymbol{g}}(\boldsymbol{\theta})$ being the asymptotic per-capita expected value of $\boldsymbol{h}$, if the distribution of $\boldsymbol{x}_{N'}$ does not change. (Intuitively, consider a sequence of actor attribute sets $\boldsymbol{x}_{N'_1}, \boldsymbol{x}_{N'_2}, \ldots$ such that $\boldsymbol{x}_{N'_i}$ is $\boldsymbol{x}_{N'_1}$ replicated i times.)

Verifying the property (A.2) requires deriving a closed form for $\boldsymbol{\mu}_{\boldsymbol{g}}(\cdot)$—at least asymptotically. Krivitsky et al. (2011, Sec. 4.3) showed this property for some dyadic-independent ERGM terms, but for dyadic-dependent ERGMs, it may not be possible to do so. In practice, this property only needs to hold in the neighborhood of the estimate $\tilde{\boldsymbol{\theta}}$, so it can be can be checked by simulating from the fitted (A.3) at a variety of network sizes with the same distributions of $\boldsymbol{x}_{N'}$, to confirm that $|N'|^{-1}\boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\boldsymbol{\theta})$ does not vary substantially in $|N'|$ for $|N'|$ large enough.

A.2. Point estimation. A model fit to network of size $|N'|$ approximating the coefficients of a model fit to network of size $|N|$ has the form

$$\Pr_{\boldsymbol{g}}(\boldsymbol{Y} = \boldsymbol{y}; \boldsymbol{x}_{N'}, \boldsymbol{\theta}) \equiv \frac{\exp\{-\log(|N'|/|N|)|\boldsymbol{y}| + \boldsymbol{\theta}^\top \boldsymbol{g}(\boldsymbol{y}, \boldsymbol{x}_{N'})\}}{\kappa_{\boldsymbol{g}}(\boldsymbol{\theta}, \boldsymbol{x}_{N'})}, \ \boldsymbol{y} \in 2^{\mathbb{Y}(N')}, \tag{A.3}$$

for $\boldsymbol{g}(\boldsymbol{y}, \boldsymbol{x}_{N'})$ estimated by $\tilde{\boldsymbol{g}}^{N'}(e_S) \equiv \tilde{\boldsymbol{g}}(e_S) \times |N'|/|N|$. Intuitively, the smaller a fraction of $|N|$ that $|N'|$ is, the more positive the offset coefficient on $|\boldsymbol{y}|$ is, forcing $\boldsymbol{\theta}$ to adjust to produce the more sparse network that N would induce. (More concretely, if, for some k, $g_k(\boldsymbol{y}) = |\boldsymbol{y}|$, its PMLE coefficient would be shifted by $\log(|N'|/|N|)$). It is still a regular exponential family, so the PMLE can be found by solving

$$\widetilde{\mathrm{sc}}^{N'}(\tilde{\boldsymbol{\theta}}^{N'}) = \tilde{\boldsymbol{g}}^{N'}(\boldsymbol{e}_S) - \boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\tilde{\boldsymbol{\theta}}^{N'}) = \boldsymbol{0},$$

where $\boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\boldsymbol{\theta})$ is the expected value of $\boldsymbol{g}(\boldsymbol{Y}, \boldsymbol{x}_{N'})$ under (A.3).

A.3. Evaluation of uncertainty. In a network process fulfilling the heuristics of Krivitsky et al. (2011), the distribution of individual measurements $\boldsymbol{h}(\boldsymbol{e}_i)$ should not be affected by the network size: the view of each individual in the network should not, for a sufficiently large network, be affected by how large the network is. Therefore, $\boldsymbol{\Sigma}_{[w,w\boldsymbol{h}]}$ in (4.4) does not depend on $|N|$, and, if (A.2) holds, the per-capita network statistics of interest should converge.

Then, provided $\boldsymbol{\lambda}_{\boldsymbol{g}}(\boldsymbol{\theta})$ itself is differentiable,

$$\lim_{|N'|\to\infty} |N'|^{-1}\nabla_{\boldsymbol{\theta}}\boldsymbol{\mu}_{\boldsymbol{g}}(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} \lim_{|N'|\to\infty} |N'|^{-1}\boldsymbol{\mu}_{\boldsymbol{g}}(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}}\boldsymbol{\lambda}_{\boldsymbol{g}}(\boldsymbol{\theta}).$$

Then, $|N|^{-1}\nabla_{\boldsymbol{\theta}}\boldsymbol{\mu}_{\boldsymbol{g}}\{\boldsymbol{\theta}_{\boldsymbol{g}}(\boldsymbol{\mu})\}$ in (4.4) can be approximated by $|N'|^{-1}\nabla_{\boldsymbol{\theta}}\boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\tilde{\boldsymbol{\theta}})$, estimated as in (4.5). This means that the asymptotic variance from (4.4) with N' in place of N,

$$\begin{aligned}
\lim_{|N'|\to\infty} \mathrm{var}_S(\tilde{\boldsymbol{\theta}}^{N'}) &= \lim_{|N'|\to\infty} \{\nabla_{\boldsymbol{\theta}}\boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\boldsymbol{\theta})\}^{-1}|N'|^2\boldsymbol{\Sigma}_{\mathrm{H}}/|S|[\{\nabla_{\boldsymbol{\theta}}^{N'}\boldsymbol{\mu}_{\boldsymbol{g}}(\boldsymbol{\theta})\}^{-1}]^{\top} \\
&= \lim_{|N'|\to\infty} \{|N'|^{-1}\nabla_{\boldsymbol{\theta}}\boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\boldsymbol{\theta})\}^{-1}\boldsymbol{\Sigma}_{\mathrm{H}}/|S|[\{|N'|^{-1}\nabla_{\boldsymbol{\theta}}\boldsymbol{\mu}_{\boldsymbol{g}}^{N'}(\boldsymbol{\theta})\}^{-1}]^{\top} \\
&= \{\nabla_{\boldsymbol{\theta}}\boldsymbol{\lambda}_{\boldsymbol{g}}(\boldsymbol{\theta})\}^{-1}\boldsymbol{\Sigma}_{\mathrm{H}}/|S|[\{\nabla_{\boldsymbol{\theta}}\boldsymbol{\lambda}_{\boldsymbol{g}}(\boldsymbol{\theta})\}^{-1}]^{\top},
\end{aligned}$$

so the variance of the estimator ceases to depend on $|N'|$ for $|N'|$ sufficiently large. This is a logical and welcome result: the variance of the estimator depends primarily on the sample size, not the population size.

A.4. Scaled estimation procedure. This leads to the following estimation procedure:

1. Construct a pseudopopulation N' that is a "scaled-down" N: i.e., the distribution of $\boldsymbol{x}_{N'}$ must be the same as $\boldsymbol{x}_N$.
2. Estimate the scaled sufficient statistic of the ERGM with $\tilde{\boldsymbol{g}}(\boldsymbol{e}_S) \times |N'|/|N|$.
3. Obtain $\tilde{\boldsymbol{\theta}}$, using MCMLE to solve $\widetilde{\mathrm{sc}}^{N'}(\tilde{\boldsymbol{\theta}}) = \boldsymbol{0}$.
4. As a byproduct of Step 3, obtain $\widetilde{\mathrm{var}}_{\boldsymbol{g}}\{\boldsymbol{g}(\boldsymbol{Y}, \boldsymbol{x}_{N'}); \tilde{\boldsymbol{\theta}}\}$.
5. Estimate $\mathrm{var}_S(\tilde{\boldsymbol{\theta}})$ as described in Section 4.2, using N' in place of N and $\boldsymbol{x}_{N'}$ in place of $\boldsymbol{x}$.
6. Simulate from the model fit for a variety of $|N'|$ to test property (A.2).

APPENDIX B: SIMULATION STUDY DETAILS AND RESULTS

In this appendix, we give more details on the simulation study and the results.

TABLE 4
Population network features

Feature	$\boldsymbol{g}(\boldsymbol{y})$	Target	Deviation[1]	$\boldsymbol{\theta}$
Total ties	$\lvert\boldsymbol{y}\rvert$	$\frac{3}{4}\lvert N\rvert = 75$k	-1.00	-10.394
Isolate count	$\sum_{i\in N} 1_{\boldsymbol{y}_i=\emptyset}$	$\frac{1}{5}\lvert N\rvert = 20$k	-1.00	1.180
Degree 1 count	$\sum_{i\in N} 1_{\lvert\boldsymbol{y}_i\rvert=1}$	$\frac{1}{2}\lvert N\rvert = 50$k	-1.00	1.555
Ties on B actors	$\sum_{(i,j)\in\boldsymbol{y}}(1_{x_{i,l}=\mathrm{B}} + 1_{x_{j,l}=\mathrm{B}})$	$1\lvert N\rvert = 100$k	0.00	0.246
Ties on C actors	$\sum_{(i,j)\in\boldsymbol{y}}(1_{x_{i,l}=\mathrm{C}} + 1_{x_{j,l}=\mathrm{C}})$	$\frac{1}{4}\lvert N\rvert = 25$k	0.00	0.000
Within-group ties	$\sum_{(i,j)\in\boldsymbol{y}} 1_{x_{i,l}=x_{j,l}}$	$\frac{1}{2}\lvert N\rvert = 50$k	0.00	1.004
Difference in $\boldsymbol{x}_{\cdot,2}$	$\sum_{(i,j)\in\boldsymbol{y}}\lvert x_{i,2} - x_{j,2}\rvert$	$\frac{1}{2}\lvert N\rvert = 50$k	$+0.06$	-0.916

[1] — Here, "Deviation" refers to the difference between the statistic of the network generated and the target value.

B.1. Study design.

B.1.1. *Simulated population network.* The population network $\boldsymbol{y}$ of size $|N| = 100{,}000$ was constructed to have the following distribution of actor attributes:

$x_{i,1}$ categorical attribute with the following composition: "A" (25%), "B" (50%), and "C" (25%); and

$x_{i,2}$ quantitative attribute, generated from $\mathrm{N}(0, 1)$ distribution.

Simulated annealing was used to find a configuration of ties such that the network statistics of interest—listed in Table 4—were as close as possible to their target values, and the differences between the generated network's statistics and the target values are shown in the same table. This network serves as the population network $\boldsymbol{y}$ in this study. (In each case, the difference is negligible, compared to the magnitude of the statistic.) An ERGM was fit to the resulting network, producing $\boldsymbol{\theta}$.

B.1.2. *Sampling design.* We considered two sample sizes, both taken without replacement: $|S| = 1{,}000$, for a sampling fraction of 1%, and $|S| = 2{,}000$, for a sampling fraction of 2%; and we considered two sampling designs: a simple random sample and a design with sampling weights that mimic sources of sampling weights that arise in applications, including oversampling of smaller subpopulations and a response rate that varied with the continuous covariate.

For each of the four combinations of sample size and weighting scheme, we drew 2,000 egocentric samples. For each of these 8,000 samples, we used the scaled estimation procedure described in Section A.4 with $|N'| \approx 1 \times |S| \approx 1{,}000$ and 2,000, $5 \times |S| \approx 5{,}000$ and 10,000, and $10 \times |S| \approx 10{,}000$ and 20,000 to estimate $\boldsymbol{\theta}$ and evaluate uncertainty.

B.1.3. *Sampling weights.* We considered two sources of unequal sampling probabilities:

- Small subpopulations can be oversampled to facilitate separate inference about them. In our simulation, we reproduce this scenario by oversampling A and C actors by a factor of 2.
- Sampling weights are also used to control for nonresponse. We emulate this by setting the response rate of actor i to be proportional to $e^{x_{i,2}}$, though for the sake of simplicity, we assume that actors are drawn from the population until the target sample size is reached.

This leads to the following sampling probabilities

$$\pi_i \propto \exp\{1_{x_{i,1}\in\{\mathrm{A,C}\}} \log(2) + x_{i,2}\}, \quad i \in N$$

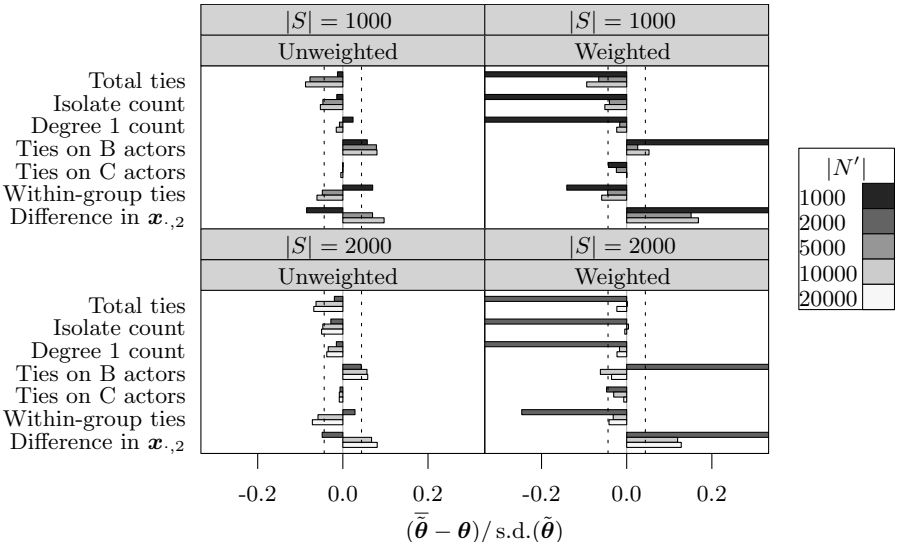
and $w_i \propto 1/\pi_i$.

B.2. Results. We summarize the biases in the point estimate for Figure 1a and compare the standard deviation of the sampling distributions of the parameter estimates to the standard errors produced by the procedure in Figure 1b. Deviations from nominal coverage are visualized in Figure 2. Numerical summaries can be found in Appendix B.3.

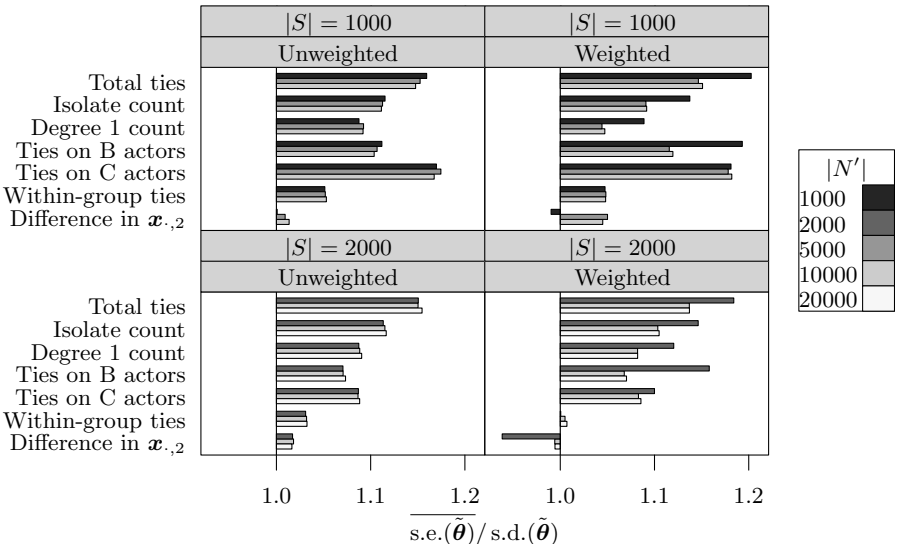
In the following section, we focus on the raw observations and patterns, and we discuss likely reasons and sources of bias in Section 8 in the main paper.

The unweighted sampling estimates display some bias, though it does not appear to have a systematic pattern as a function of $|N'|$ or in the model term. For $|S| = 1{,}000$, none of the estimated biases are greater than 10% of the standard deviation under repeated sampling, which is to say that bias accounts for less than 1% of the mean squared error (MSE) of the estimator. For $|S| = 2{,}000$ they are even smaller relative to their standard deviation (which is, itself, about $\sqrt{2}$ times smaller).

The weighted sampling estimators are, as one would expect, highly biased for smaller $|N'|$. For the largest $|N'|$, the bias of the most biased parameter estimate (Difference in $\boldsymbol{x}_{\cdot,2}$) is less than 20% of the standard deviation under repeated sampling (i.e., about 4% of the total MSE), even for $|S| = 1{,}000$. A possible reason why this particular estimate is the most biased is that egos with small $\boldsymbol{x}_{i,2}$ are (by design) severely undersampled, which means that there will exist many samples where the full range of $\boldsymbol{x}_{\cdot,2}$ is not represented. This is likely to be less problematic in real-world applications like the analysis in Section 7, where continuous covariates (like age) have an explicit range of interest. As expected, estimators under $|S| = 2{,}000$ exhibit uniformly smaller bias, even as a fraction of the smaller standard deviation.



(a) Bias in point estimates



(b) Bias in standard errors

Fig 1: Simulated bias in the point estimates and standard errors: the point estimates, normalized by s.d.($\tilde{\boldsymbol{\theta}}$). Dashed lines are positioned at $\pm 1.96/\sqrt{2000}$, so about 95% of the simulated biases should fall into these intervals if their true mean is 0. Some of the biases are truncated to preserve detail.

Overall, the standard errors under unweighted sampling appear to be conservative, overestimating the simulated standard deviation under repeated sampling by between 1% and 20% in some cases, and there is evidence of them becoming more accurate as the sample size increases.

This positive bias in standard errors may be a consequence of estimating the distribution of $\boldsymbol{x}_N$ from $\boldsymbol{x}_S$: somewhat counterintuitively, it may reduce the actual variance of $\tilde{\boldsymbol{\theta}}$ in the presence of homophily, because those samples that happen to contain, say, an excess of members of Group B will also contain an excess of ties incident on members of Group B, which is consistent with ERGM behavior for changing composition (Krivitsky et al., 2011).

The resulting Wald confidence interval coverage, summarized in Figure 2, is consistent with the above observations: in almost all terms, the intervals are somewhat conservative for both unweighted and weighted sampling (given sufficient $|N'|$), likely a consequence of the variance being overestimated. The coverage does appear to improve with the sample size.

We also find that for $|S| = 1{,}000$, while most parameter estimates' sampling distributions were statistically indistinguishable from normal (based on 2,000 simulated realizations each), the parameters corresponding to total number of ties and number of ties incident on actors in Group B show slight deviations from normality in both unweighted and weighted simulations. (Shapiro–Wilk P-val. < 0.01 for each.) The former term's parameter estimates exhibit negative skewness while the latter term's exhibit positive skewness. This may be because their corresponding statistics are fairly strongly negatively correlated with each other (because, with B being the largest group, and there being positive within-group homophily, 82% of the ties in $\boldsymbol{y}$ involve an actor in Group B). This strong correlation may be slowing down the rate at which their joint distribution asymptotes, further exacerbated by actors in Group B being undersampled. For $|S| = 2{,}000$, none are significantly non-normal.

These results are encouraging, in that even with a fairly modest sample size, and in the presence of fairly heavy weighting, the confidence intervals are reasonable, provided $|N'|$ is sufficiently large. In particular, additional error due to the distribution of $\boldsymbol{x}_N$ being inferred from the sample does not appear to invalidate them.

B.3. Simulation study summary tables. The following tables give numerical summaries of the simulation studies. $\overline{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$ is the bias of the point estimates and s.d.$(\tilde{\boldsymbol{\theta}})$ is their simulated standard deviation, both obtained based on 2,000 replications of egocentric sampling and estimation; and $\overline{\text{s.e.}(\tilde{\boldsymbol{\theta}})}$ is the mean of the standard errors calculated from (4.5) for each

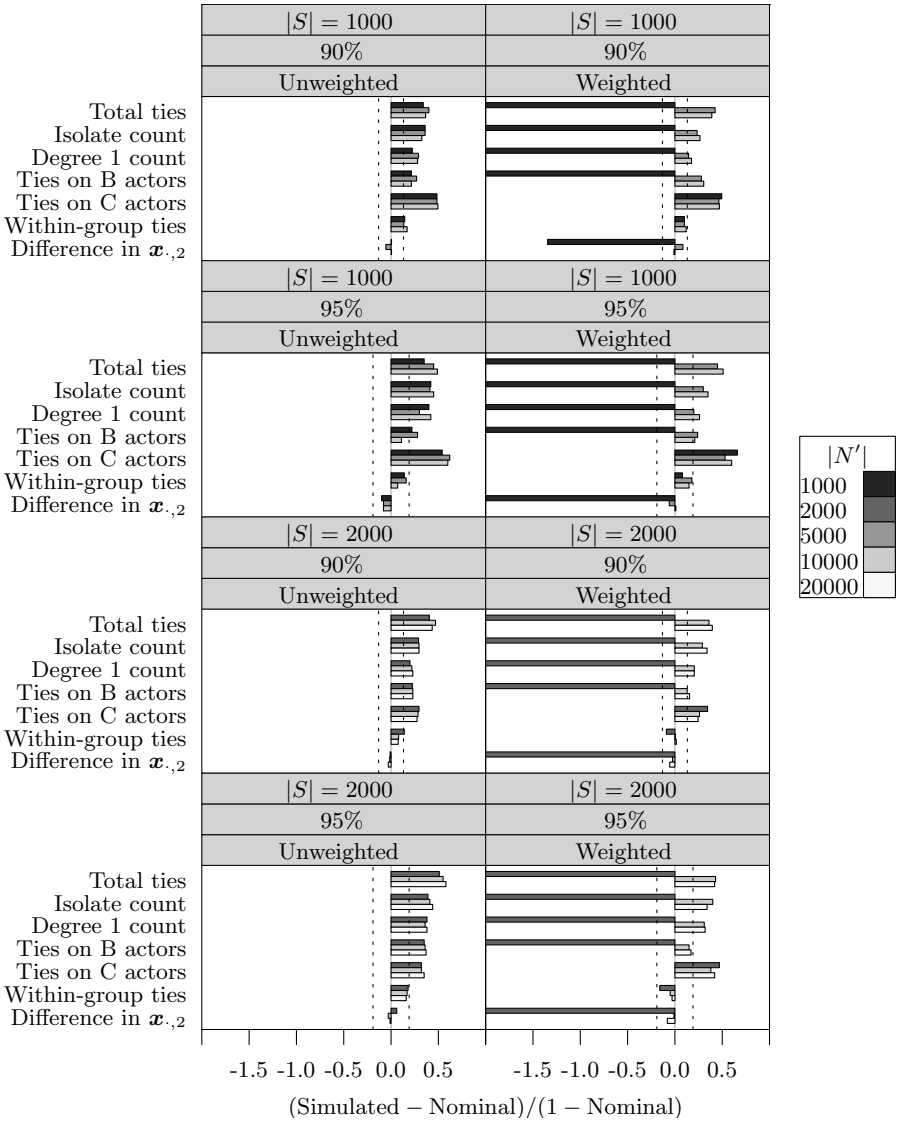


Fig 2: Coverage simulation results: the difference between the simulated and the nominal coverage is given, relative to the nominal probability of a miss. (Note that this quantity cannot be greater than 1.) Dashed lines are positioned at $\pm 1.96\sqrt{\mathrm{CL}(1-\mathrm{CL})/2000}/(1-\mathrm{CL})$, so about 95% of the simulated coverages should fall into these intervals if they do, on average, equal to nominal. Some of the coverages are truncated to preserve detail.

replication. Coverages for 90%, 95%, and 99% Wald confidence intervals are also given.

B.3.1. $|S| = 1{,}000$, $|N'| = 1{,}000$.

Summaries:

		Unweighted			Weighted		
	$\boldsymbol{\theta}$	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)
Total ties	−10.394	−0.002	0.168	0.195	−0.789	0.273	0.328
Isolate count	1.180	−0.003	0.183	0.204	−0.569	0.280	0.319
Degree 1 count	1.555	0.003	0.118	0.129	−0.258	0.156	0.170
Ties on B actors	0.246	0.004	0.077	0.085	0.398	0.117	0.140
Ties on C actors	0.000	0.000	0.082	0.096	−0.004	0.101	0.119
Within-group ties	1.004	0.004	0.063	0.066	−0.009	0.065	0.069
Difference in $\boldsymbol{x}_{\cdot,2}$	−0.916	−0.004	0.051	0.051	0.052	0.060	0.059

Coverage:

	Unweighted			Weighted		
	90	95	99	90	95	99
Total ties	93.4	96.8	99.6	10.1	22.0	63.6
Isolate count	93.6	97.1	99.6	43.1	60.0	87.2
Degree 1 count	92.2	97.0	99.4	55.3	68.0	88.6
Ties on B actors	92.2	96.1	99.1	1.7	5.7	29.8
Ties on C actors	94.8	97.7	99.7	95.0	98.3	99.8
Within-group ties	91.5	95.7	99.2	91.0	95.4	99.2
Difference in $\boldsymbol{x}_{\cdot,2}$	90.0	94.5	98.9	76.5	84.7	94.3

B.3.2. $|S| = 1{,}000$, $|N'| = 5{,}000$.

Summaries:

		Unweighted			Weighted		
	$\boldsymbol{\theta}$	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)
Total ties	−10.394	−0.013	0.169	0.195	−0.012	0.183	0.210
Isolate count	1.180	−0.009	0.183	0.204	−0.009	0.215	0.234
Degree 1 count	1.555	−0.001	0.118	0.129	−0.002	0.143	0.149
Ties on B actors	0.246	0.006	0.077	0.085	0.002	0.076	0.085
Ties on C actors	0.000	0.000	0.082	0.096	−0.002	0.076	0.090
Within-group ties	1.004	−0.003	0.063	0.066	−0.003	0.065	0.068
Difference in $\boldsymbol{x}_{\cdot,2}$	−0.916	0.003	0.050	0.050	0.008	0.056	0.059

Coverage:

	Unweighted			Weighted		
	90	95	99	90	95	99
Total ties	94.0	97.2	99.5	94.2	97.2	99.2
Isolate count	93.6	97.0	99.6	92.3	96.5	99.3
Degree 1 count	92.9	96.5	99.4	91.5	96.0	99.2
Ties on B actors	92.7	96.4	99.3	92.8	96.2	98.9
Ties on C actors	94.8	98.1	99.8	94.7	97.7	99.6
Within-group ties	91.3	95.8	99.2	91.0	95.9	99.4
Difference in $\boldsymbol{x}_{\cdot,2}$	89.5	94.6	98.8	90.8	94.7	99.2

B.3.3. $|S| = 1{,}000$, $|N'| = 10{,}000$.

Summaries:

		Unweighted			Weighted		
	$\boldsymbol{\theta}$	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)
Total ties	−10.394	−0.015	0.169	0.194	−0.017	0.183	0.210
Isolate count	1.180	−0.010	0.183	0.204	−0.011	0.215	0.235
Degree 1 count	1.555	−0.002	0.118	0.129	−0.003	0.143	0.149
Ties on B actors	0.246	0.006	0.077	0.084	0.004	0.076	0.085
Ties on C actors	0.000	0.000	0.082	0.096	0.000	0.076	0.090
Within-group ties	1.004	−0.004	0.063	0.066	−0.004	0.065	0.068
Difference in $\boldsymbol{x}_{\cdot,2}$	−0.916	0.005	0.050	0.050	0.009	0.056	0.058

Coverage:

	Unweighted			Weighted		
	90	95	99	90	95	99
Total ties	93.7	97.5	99.5	93.9	97.5	99.2
Isolate count	93.2	97.2	99.5	92.7	96.8	99.4
Degree 1 count	92.8	97.1	99.4	91.8	96.3	99.3
Ties on B actors	92.2	95.5	99.2	93.0	96.0	99.1
Ties on C actors	95.0	98.0	99.8	94.7	98.0	99.8
Within-group ties	91.7	95.3	99.2	91.2	95.8	99.2
Difference in $\boldsymbol{x}_{\cdot,2}$	90.0	94.6	98.6	89.9	95.0	99.1

B.3.4. $|S| = 2{,}000$, $|N'| = 2{,}000$.

Summaries:

		Unweighted			Weighted		
	$\boldsymbol{\theta}$	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)
Total ties	−10.394	−0.002	0.116	0.133	−0.773	0.188	0.223
Isolate count	1.180	−0.004	0.126	0.141	−0.555	0.190	0.218
Degree 1 count	1.555	−0.001	0.082	0.089	−0.258	0.104	0.117
Ties on B actors	0.246	0.002	0.054	0.057	0.390	0.082	0.095
Ties on C actors	0.000	0.000	0.060	0.065	−0.003	0.073	0.081
Within-group ties	1.004	0.001	0.043	0.045	−0.011	0.046	0.046
Difference in $\boldsymbol{x}_{\cdot,2}$	−0.916	−0.002	0.034	0.034	0.052	0.043	0.040

Coverage:

	Unweighted			Weighted		
	90	95	99	90	95	99
Total ties	94.0	97.5	99.6	0.2	0.8	6.6
Isolate count	92.9	97.0	99.5	11.3	22.3	52.8
Degree 1 count	92.0	96.9	99.6	26.3	39.0	66.3
Ties on B actors	92.2	96.8	99.3	0.0	0.0	0.4
Ties on C actors	93.0	96.6	99.4	93.5	97.4	99.7
Within-group ties	91.4	95.9	99.4	89.1	94.2	99.0
Difference in $\boldsymbol{x}_{\cdot,2}$	89.9	95.3	99.2	61.8	73.1	87.8

B.3.5. $|S| = 2{,}000$, $|N'| = 10{,}000$.

Summaries:

		Unweighted			Weighted		
	$\boldsymbol{\theta}$	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)	$\bar{\tilde{\boldsymbol{\theta}}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)
Total ties	−10.394	−0.007	0.116	0.133	0.000	0.126	0.143
Isolate count	1.180	−0.006	0.126	0.141	0.001	0.146	0.161
Degree 1 count	1.555	−0.003	0.082	0.089	−0.002	0.095	0.103
Ties on B actors	0.246	0.003	0.054	0.057	−0.003	0.054	0.057
Ties on C actors	0.000	0.000	0.060	0.065	−0.002	0.056	0.060
Within-group ties	1.004	−0.003	0.043	0.045	−0.001	0.046	0.046
Difference in $\boldsymbol{x}_{\cdot,2}$	−0.916	0.002	0.034	0.034	0.005	0.040	0.040

Coverage:

	Unweighted			Weighted		
	90	95	99	90	95	99
Total ties	94.7	97.8	99.6	93.6	97.2	99.4
Isolate count	93.0	97.0	99.6	92.9	97.0	99.5
Degree 1 count	92.2	96.8	99.6	92.0	96.5	99.3
Ties on B actors	92.3	96.8	99.4	91.3	95.8	99.1
Ties on C actors	92.8	96.6	99.5	92.6	96.9	99.6
Within-group ties	90.8	95.9	99.4	90.0	94.8	99.0
Difference in $\boldsymbol{x}_{\cdot,2}$	89.8	94.8	99.2	89.8	95.0	99.1

B.3.6. $|S| = 2{,}000$, $|N'| = 20{,}000$.

Summaries:

		Unweighted			Weighted		
	$\boldsymbol{\theta}$	$\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)	$\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}$	s.d.($\tilde{\boldsymbol{\theta}}$)	s.e.($\tilde{\boldsymbol{\theta}}$)
Total ties	−10.394	−0.008	0.115	0.133	−0.003	0.126	0.143
Isolate count	1.180	−0.006	0.126	0.141	−0.001	0.146	0.161
Degree 1 count	1.555	−0.003	0.082	0.089	−0.002	0.095	0.103
Ties on B actors	0.246	0.003	0.054	0.058	−0.002	0.053	0.057
Ties on C actors	0.000	−0.001	0.060	0.065	0.000	0.056	0.060
Within-group ties	1.004	−0.003	0.043	0.045	−0.002	0.046	0.046
Difference in $\boldsymbol{x}_{\cdot,2}$	−0.916	0.003	0.034	0.034	0.005	0.040	0.040

Coverage:

	Unweighted			Weighted		
	90	95	99	90	95	99
Total ties	94.3	97.9	99.6	94.0	97.1	99.4
Isolate count	93.0	97.2	99.7	93.4	96.7	99.4
Degree 1 count	92.3	96.9	99.6	92.0	96.6	99.3
Ties on B actors	92.3	96.9	99.2	91.5	95.9	99.1
Ties on C actors	92.8	96.8	99.4	92.5	97.1	99.6
Within-group ties	90.8	95.8	99.5	90.1	94.8	99.0
Difference in $\boldsymbol{x}_{\cdot,2}$	89.7	95.0	99.1	89.5	94.6	98.9

APPENDIX C: AUXILIARY RESULTS FOR THE APPLICATION

In this appendix, we report auxiliary results to our analysis in Section 7. In particular, we discuss the heuristic for our choice of $|N'|$, confirm that our results remain after controlling for age, and verify assumption (A.2) as it pertains to the models fit.

C.1. Selecting $|N'|$ for the analysis. Our choice of $|N'|$ is driven by the data including sampling weights: inference requires that the N' be as representative of N as possible, and, as we show in the Appendix B, insufficient $|N'|$ can significantly bias estimation. We cannot compare the two situations directly, but, heuristically, the weights are somewhat more varied in the NHSLS data than in the weighted simulation study: in the simulation study's samples of 2,000, $w_{\max}/w_{\min}$ averaged 9.1 for $|S| = 1{,}000$ and 9.9 for $|S| = 2{,}000$, and in the NHSLS study (after excluding respondents with missing data), this ratio is somewhat higher, 15.3 (albeit at a greater sample size). Another metric is the amount by which the variance of the sample mean would be inflated due to unequal weighting, relative to an SRS, which equals to $|N|^{-1}\sum_{i=1}^{|N|} w_i^2/\bar{w}^2$. This is 1.18 for the simulation study and 1.34 for the NHSLS data. $|N'| \approx 5|S|$ appear to be adequate for the simulated data, though $|N'| \approx 10|S|$ produces a noticeable improvement, so we select, conservatively, $|N'| \approx 45{,}000 \approx 13.4 \times |S|$. The respondent with the smallest sampling weight represents about $w_i/w_\cdot = 7.28 \times 10^{-5}$ of the

population, so she is represented in N' about three times.

C.2. Detailed goodness of fit. In this section, we report goodness of fit broken down by race and sex. Figure 3 gives a plot analogous to those for Figure 2a for each combination of these two factors. As expected, *Model 3* still clearly has the best fit, though some specific groups have slight deviations. In particular, for the "Other" Males, the model overestimates slightly the number of isolate actors and underestimates the number of concurrent; and for Black Males, it slightly underestimates the number of individuals with two concurrent partners but overestimates the number with three by about the same amount.

Because these are some of the smaller groups with some of the highest rates of nonresponse, this may be a consequence of the estimates themselves being uncertain. In any case, the overall level of concurrency in all groups except "Other" Males is well modelled.

C.3. Controlling for age. In this section, we report a model fit that, in addition to representing mixing by race and monogamy, also incorporates age effects.

In this, we follow the analysis of these data by Krivitsky et al. (2011), modeling the effects of age semiparametrically. As predictors, we consider the age of the actor, the square root of age, the age difference and squared difference in a potential partnership, and the difference and the squared difference of the square roots of ages. To improve numeric conditioning of the model, we perform an affine transformation on the ages, shifting and scaling them into a $[-1/2, +1/2]$ interval: $x'_{i,\text{age}} = (x_{i,\text{age}} - 18)/(60 - 18) - 1/2$. This change merely scales the coefficient and changes the baseline coefficients (number of ties, by sex), without changing the family of distributions being modeled. For the square root of age effects, the corresponding transformation is

$$x'_{i,\sqrt{\text{age}}} = \sqrt{\frac{x_{i,\text{age}} - 18}{60 - 18}} - \frac{1}{2}.$$

The use of the square root and linear effect, rather than linear and quadratic, is motivated by the notion that the effect of a one-year difference will be greater for younger actors than older: going from 20 to 21 is likely to have a greater effect than going from 50 to 51.

The results for *Model 3* with age, along with the estimates for *Model 3* itself for comparison, are given in Table 5. The most important aspect of this result is that the coefficients estimated for the Monogamy model have not changed qualitatively after age effects are controlled for: our results in

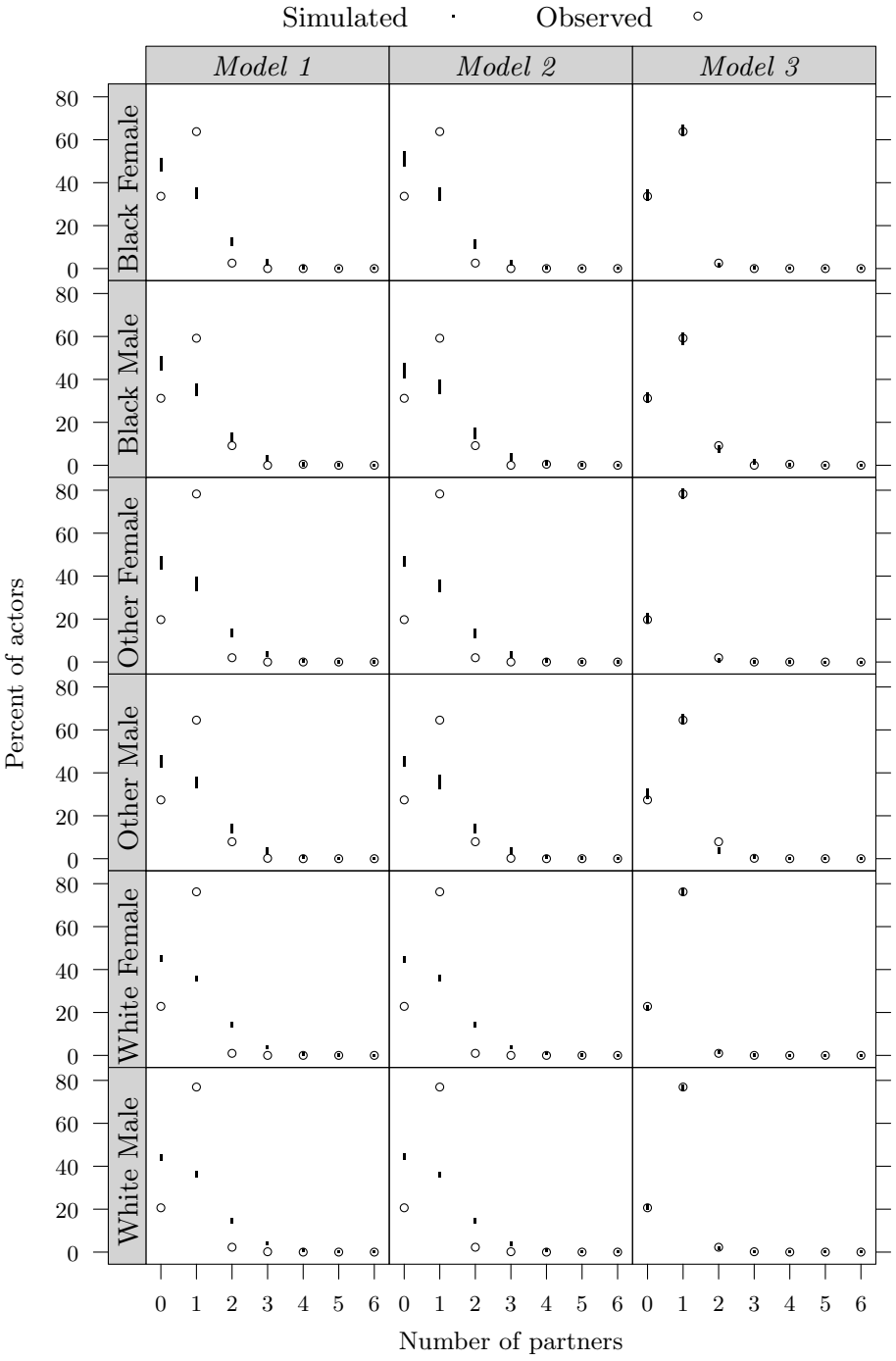


Fig 3: Goodness of fit for degree distribution, for each model, broken down by sex and race. For each of the six combinations, the degree frequencies (dot plot) for 100 realizations is compared to those observed in the data.

TABLE 5

Coefficients and significance for the Model 3 *(main, mixing, and monogamy effects) and a model that also incorporates age effects. Coefficients reported are in the presence of an edge count offset of* $-\log(44859) = -10.71$.

	Main + Mix. + Monog.	+ Age
Actor activity by sex		
Female	$-1.88\ (0.31)^{***}$	$-1.78\ (0.40)^{***}$
Male	$-1.18\ (0.25)^{***}$	$-1.08\ (0.35)^{**}$
Same-sex partnership	$-4.52\ (0.21)^{***}$	$-4.13\ (0.21)^{***}$
Actor activity by race		
White	0 (baseline)	
Black	$-0.30\ (0.38)$	$-0.35\ (0.36)$
Other	$0.93\ (0.42)^{*}$	$0.87\ (0.43)^{*}$
Race homophily by race		
Black	$5.15\ (0.38)^{***}$	$5.16\ (0.38)^{***}$
Other	$2.04\ (0.35)^{***}$	$2.09\ (0.39)^{***}$
White	$2.32\ (0.36)^{***}$	$2.31\ (0.39)^{***}$
Monogamy by sex and race		
Black female	$1.80\ (0.47)^{***}$	$1.94\ (0.50)^{***}$
Other female	$2.51\ (0.67)^{***}$	$2.56\ (0.69)^{***}$
White female	$2.25\ (0.31)^{***}$	$2.36\ (0.32)^{***}$
Black male	$0.99\ (0.24)^{***}$	$1.11\ (0.28)^{***}$
Other male	$1.40\ (0.31)^{***}$	$1.54\ (0.33)^{***}$
White male	$2.16\ (0.25)^{***}$	$2.30\ (0.25)^{***}$
Age effects		
$\sqrt{\text{age}}$ effect		$3.29\ (1.35)^{*}$
age effect		$-2.73\ (1.15)^{*}$
Age difference effects		
Difference in $\sqrt{\text{age}}$		$-8.07\ (2.20)^{***}$
Difference in age		$-6.03\ (1.92)^{**}$
Squared difference in $\sqrt{\text{age}}$		$3.22\ (3.73)$
Squared difference in age		$2.32\ (2.80)$
Older-male-younger-female		$0.93\ (0.05)^{***}$

Significance levels: $0.05 \geq {}^{*} > 0.01 \geq {}^{**} > 0.001 \geq {}^{***}$

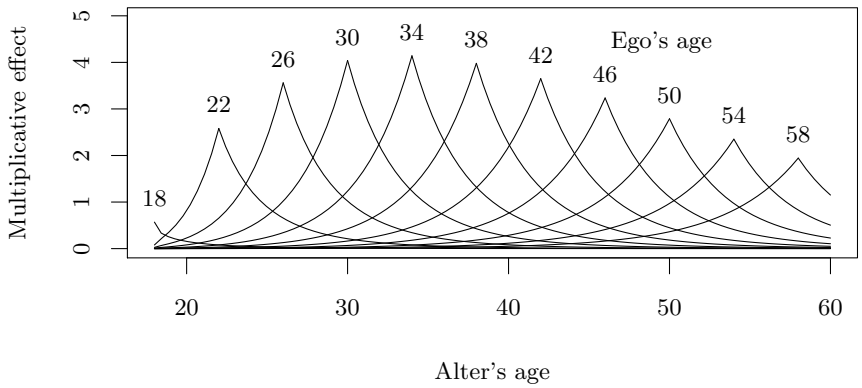


Fig 4: Estimated multiplicative effects of the age of ego and age of alter (ignoring age–sex interaction) on the odds of a tie. Note that the peak of each ego age curve represents the overall propensity for egos of that age to have ties.

Section 7 are robust to age effects. (We also performed the degree distribution and component size simulations for the age model. Those were not qualitatively different either.)

Age effects themselves can be interpreted as well; we provide a visualization of the overall estimated age effect in Figure 4. Age difference effects are particularly significant. There is also strong evidence for the tendency of the male to be older than the female in a heterosexual partnership.

C.4. Simulation results to verify Assumption (A.2)**.** We test the assumption (A.2) by simulating from the most complex model fit to obtain 10,000 realizations (with some serial dependence) of $\boldsymbol{g}(\boldsymbol{Y}, \boldsymbol{x}_{N''})/|N''|$ with $|N''| \approx 90{,}000$ with offsets adjusted appropriately. If the assumption is violated, we would expect them to be different, on average, from the observed $\tilde{\boldsymbol{h}}(\boldsymbol{e}_S)$.

We report the simulation results in Table 6. The differences between the observed values and those simulated for $|N'| \approx 90,000$ are statistically significant in a few cases—as they would inevitably be, given a sufficient simulation size, but they are not practically so: the statistics with the greatest relative difference between $|N'| \approx 45,000$ and $|N'| \approx 90,000$ are ones with the smallest counts and effective numbers of observations, so one might expect them to asymptote more slowly; even among them, the greatest one has 0.120% difference.

TABLE 6

Difference between the observed per-capita statistics (denoted $\tilde{\boldsymbol{h}}(\boldsymbol{e}_S)$) and the per-capita moments of the sufficient statistics simulated from a network with $|N''| \approx 90{,}000$ using coefficients obtained with $|N'| \approx 45{,}000$ (denoted $\boldsymbol{\mu}_{\boldsymbol{g}}^{N''}(\tilde{\boldsymbol{\theta}}^{N'})/|N''|$). The differences have been scaled by 10^4 for readability, and the simulation's standard errors are adjusted for autocorrelation. Effective Sample Sizes (ESS) are also given. R (R Core Team, 2013) package `coda` *(Plummer et al., 2006) was used to evaluate the latter.*

Term	Observed $\tilde{\boldsymbol{h}}(\boldsymbol{e}_S)$	Simulated $\left\{\frac{\boldsymbol{\mu}_{\boldsymbol{g}}^{N''}(\tilde{\boldsymbol{\theta}}^{N'})}{\lvert N''\rvert} - \tilde{\boldsymbol{h}}(\boldsymbol{e}_S)\right\} \times 10^4$ (ESS, s.e.)	$\frac{\text{Diff.}}{\tilde{\boldsymbol{h}}(\boldsymbol{e}_S)}$
Actor activity by sex			
Female	0.396	0.085 (1366, 0.229)	0.002%
Male	0.399	0.033 (1429, 0.228)	0.001%
Same-sex partnership	0.005	−0.056 (4336, 0.034)	−0.120%
Actor activity by race (White as baseline)			
Black	0.087	−0.287 (564, 0.306)	−0.033%
Other	0.102	0.242 (777, 0.231)	0.024%
Race homophily by race			
Black	0.040	−0.113 (508, 0.166)	−0.028%
Other	0.038	0.136 (411, 0.184)	0.036%
White	0.288	0.025 (1421, 0.186)	0.001%
Monogamy by sex and race			
Black Female	0.042	−0.148 (735, 0.136)	−0.035%
Other Female	0.052	0.046 (1025, 0.107)	0.009%
White Female	0.284	0.358 (1749, 0.177)*	0.013%
Black Male	0.031	−0.268 (776, 0.132)*	−0.086%
Other Male	0.041	−0.189 (971, 0.122)	−0.046%
White Male	0.290	−0.053 (1979, 0.168)	−0.002%
Age effects			
$\sqrt{\text{age}}$ effect	0.104	0.142 (1285, 0.120)	†
age effect	−0.046	0.069 (1294, 0.125)	†
Age difference effects			
Difference in $\sqrt{\text{age}}$	0.028	0.130 (1135, 0.049)**	0.046%
Difference in age	0.034	0.113 (1114, 0.058)	0.034%
Squared difference in $\sqrt{\text{age}}$	0.004	0.023 (1850, 0.013)	0.055%
Squared difference in age	0.006	0.011 (1787, 0.017)	0.019%
Older-male-younger-female	0.242	0.713 (452, 0.504)	0.029%

Significance levels: $0.05 \geq {}^{\star} > 0.01 \geq {}^{\star\star} > 0.001 \geq {}^{\star\star\star}$

† — Percent differences are not meaningful for statistics that are not counts or sums of nonnegative quantities.

REFERENCES

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., and Morris, M. (2008). `ergm`: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *J. Stat. Softw.* **24** 1–29.

Krivitsky, P. N., Handcock, M. S., and Morris, M. (2011). Adjusting for Network Size and Composition Effects in Exponential-Family Random Graph Models. *Stat. Methodol.* **8** 319–339. .

Krivitsky, P. N. and Morris, M. (2016). Inference for Social Network Models from Egocentrically-Sampled Data, with Application to Understanding Persistent Racial Disparities in HIV Prevalence in the US. *Ann. Appl. Stat.* **To appear**.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). `CODA`: Convergence Diagnosis and Output Analysis for MCMC. *R News* **6** 7–11.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.