APPENDIX 1: DETAILS ON THE MATCHING ALGORITHM AND THE EGO-NETWORK OVERLAP

The matching algorithm to detect ego-network overlap is based on first and last names, sex, ethnicity, and age of individuals in the data (see Section 3.2). For anonymity requirements that are common in personal network research, alter names were in the format "F Last", with "F" being the initial of the first name and "Last" being the entire last name; however, entire first and last names were available for egos. While matching errors are possible with the proposed algorithm, we are confident that the procedure achieves a reasonable degree of accuracy by taking into consideration both names and multiple sociodemographic attributes. Very similar methods, matching individual names and sociodemographics, have been successfully used in existing network studies (e.g., Mouw and Verdery 2012; Mouw et al. 2014; Merli et al. 2016). Mouw and Verdery (2012) adopted a similar procedure on data from a similar population (Mexican immigrants in the US), and assessed that results were accurate after evaluating matches in the first 1,000 nominations with the help of community members. These attribute matching techniques are a feasible and practical approach to address two issues that are typical of personal network research: (1) there is normally no feasible mechanism to detect potential overlap during data collection, because the data may be collected simultaneously from different respondents, at different locations, and on paper or off-line electronic devices; (2) complete names and identifying information may not be collected on respondents and their social contacts due to research ethics requirements (McCarty et al. 2019).

In the alter-alter overlap data, 61% of the 1,200 nominated alters are co-mentioned by multiple respondents. The result is a two-mode whole network of 60 egos connected to

710 unique alters, shown in Figure 4a. In the ego-alter overlap, 37% of the 1,200 initially nominated alters are matched to another alter, and 6% of them are matched to a respondent. Overall, 70% of respondents (42 out of 60) *nominate* other respondents in their personal network; and 42% of the respondents (25 out of 60) *are nominated* by other respondents. The result is a one-mode network with 1,200 directed edges and 745 actors (Figure 4b), including 25 actors (3% of the total) who both send and receive ties (they occupy both ego and alter positions), 35 actors who only send ties (they only appear as egos), and 685 actors who only receive ties (they only appear as alters).

<<Figure 4 about here >>

APPENDIX 2: ESTIMATION METHOD AND DIAGNOSTICS

All models in this paper are fitted with the *MCMCglmm* R package for MCMC estimation of generalized linear mixed models (Hadfield 2010). For fixed effect coefficients, a multivariate normal prior distribution is adopted with zero mean vector and a diagonal variance matrix with high variances (=1e+10). The variance-covariance matrices of the random effects are assigned a weakly informative inverse-Wishart prior distribution with a low degree of belief parameter (=.002).

Results for the variance components models (M1a-M5a) are extracted from single MCMC chains of 1,010,000 iterations, with a burn-in period of 10,000 iterations and a thinning interval of 1,000, resulting in a sample of 1,000 iterations. For the full models (M1b-M4b), single MCMC chains of 2,020,000 iterations were used, with burn-in periods of 20,000 iterations and thinning intervals of 500 iterations, resulting in a sample size of 4,000. For all chains we obtained Geweke's and Heidelberger and Welch's convergence diagnostics (results available from the authors) using functions from the *coda* R package (Plummer et al. 2006).

The Heidelberger and Welch's diagnostics are used to conduct two consecutive convergence tests. The first tests the null hypothesis that the chain values come from a stationary distribution; the second, "half-width" test checks that the ratio between half the width of the 95% confidence interval for the parameter's posterior mean, and the mean itself, is sufficiently low. All chains for all model parameters passed the Geweke's test, indicating chain convergence. All chains also passed the first Heidelberger and Welch's test. The second Heidelberger and Welch's test was not passed by the chain for one fixed effect (Count of family members) for models M3b and M4b; and by the chains for two fixed effects (Count of family members and Average alter age) for model M5b. To assess how critical the convergence of those parameter estimations was for final results, we performed sensitivity tests by removing the relevant fixed effects from M3b-M5b and re-fitting the models. Results (available from the authors) indicate that the removal of the problematic fixed effects does not substantially alter estimates for all other parameters and does not change the substantive conclusions drawn from the analysis.