

ONLINE SUPPLEMENTARY MATERIALS

Supplementary Methods & Materials

Quantitative Measures

Proximity of Home Addresses. To calculate distances between students' residences, we entered the addresses into Open Street Map [59] to obtain latitude and longitude coordinates. The latitude longitude data were imported into QGIS software version 3.16 [60] for geospatial analysis. Route-based distance between points was calculated using a shortest path analysis through the origin-destination matrix function of the QNEAT3 plug-in [61]. Results from the path analyses were then converted from meters to kilometers. We dichotomized the values in the matrix at 10 kilometers (0=*greater than 10 kilometers*, 1=*less than or equal to 10 kilometers*) as a measure of proximity.

Quantitative Data Analysis

Curved Exponential Family (CEF) Models for Correlates of Friendship Ties

Curved Exponential Family Models (CEFs, [63]) are a type of Exponential Random Graph Model (ERGM). ERGMs estimate the log odds of a tie between two network members using the following formula (equation 1):

$$P(\mathbf{Y} = \mathbf{y} \mid \theta) = \frac{1}{k} \exp\{\sum_A \theta_A z_A(\mathbf{y})\} \quad (1)$$

where \mathbf{y} is the observed network, k is a normalizing constant to make the expression a proper probability distribution, A is a network configuration consisting of ties and attributes, $z_A(\mathbf{y})$ is the count of the number of configurations of type A , and θ_A is the parameter for the configuration of type A [53]. Significant parameter estimates indicate that there are more (for positive parameter estimates) or fewer (for negative parameter estimates) configurations in the observed network than would be expected by chance given the rest of the model.

Geometric Weighting and Decay Parameters. In ERGMs, some structural effects can be estimated using a geometrically weighted count. Without geometrical weighting, each count of the structural effect in the observed network has an equal weight in the estimation. Geometrical weighting creates a diminished effect for additional counts of the structural effect. For example, if two people have a shared friend, they may be more likely to be friends themselves (i.e., a transitivity effect). This effect would be increased by having additional friends in common. In geometrical weighting, for each additional friend in common, the strength of the effect would get smaller. This diminishing return is quantified by a scaling, or decay, parameter. Not only does geometrical weighting help with the analyses, but also more importantly it reflects social behavior in the real world.

The magnitude of the decay parameter determines the amount of diminishing returns for additional structural effects. In CEFs, the decay parameter is estimated within the model [63]. In our CEF model, the effects of popularity spread, activity spread, transitive shared partners (i.e., transitivity), and two-paths. Therefore, each of these structural effects will have an estimate for the structural effect and for its decay parameter.

Goodness of Fit Plots
Preliminary Model (Figures S1 – S4)

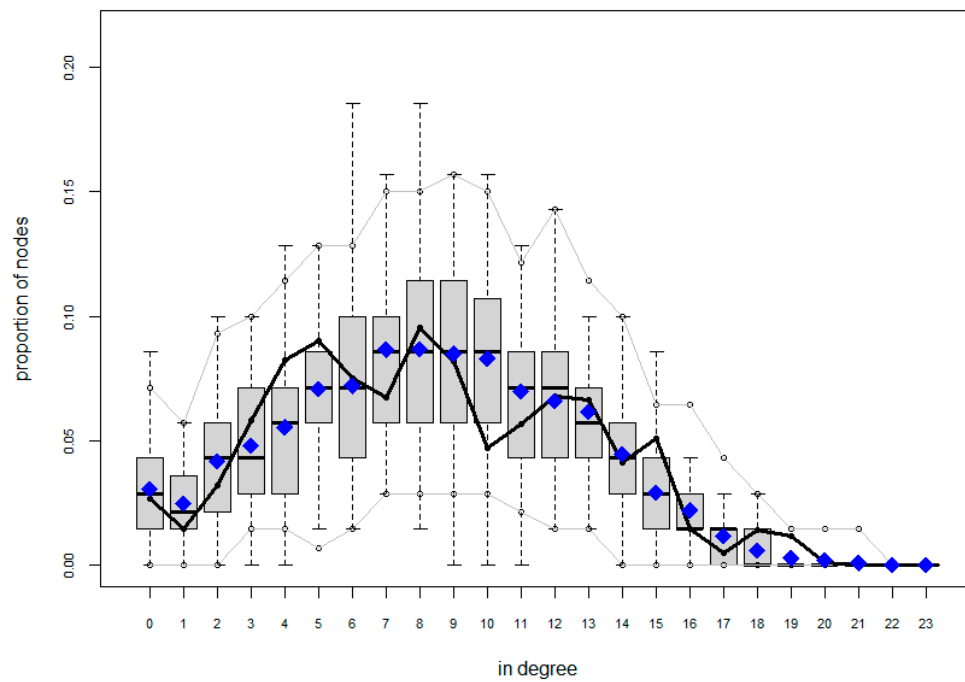


Figure S1. Goodness of fit plot for indegree distribution, preliminary model.

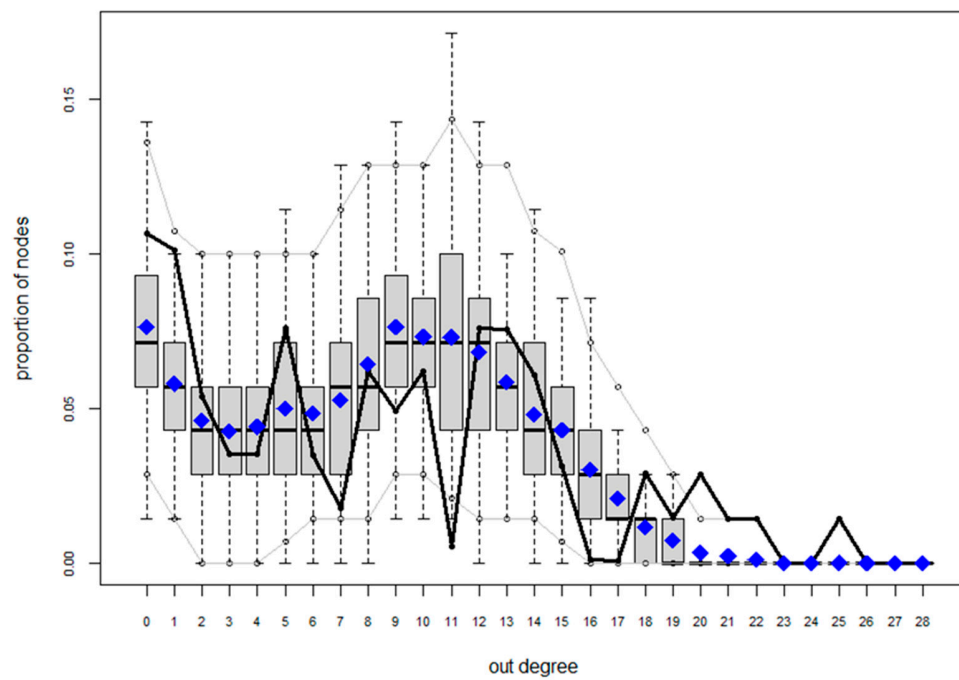


Figure S2. Goodness of fit plot for outdegree distribution, preliminary model.

Goodness of Fit Plots
Preliminary Model (continued)

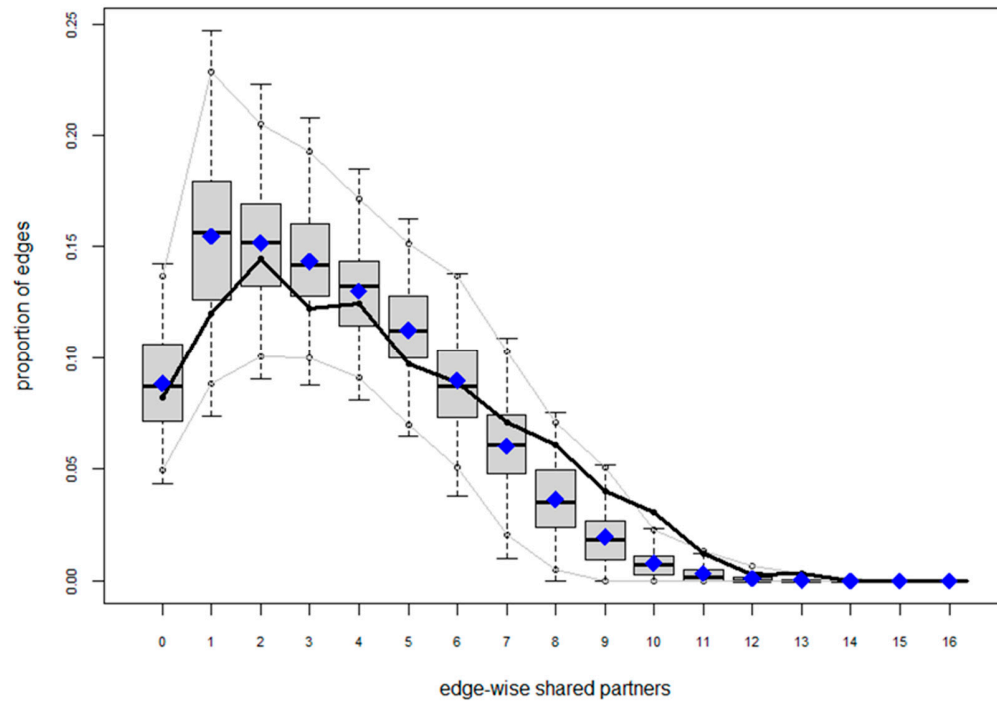


Figure S3. Goodness of fit plot for edge-wise shared partners distribution, preliminary model.

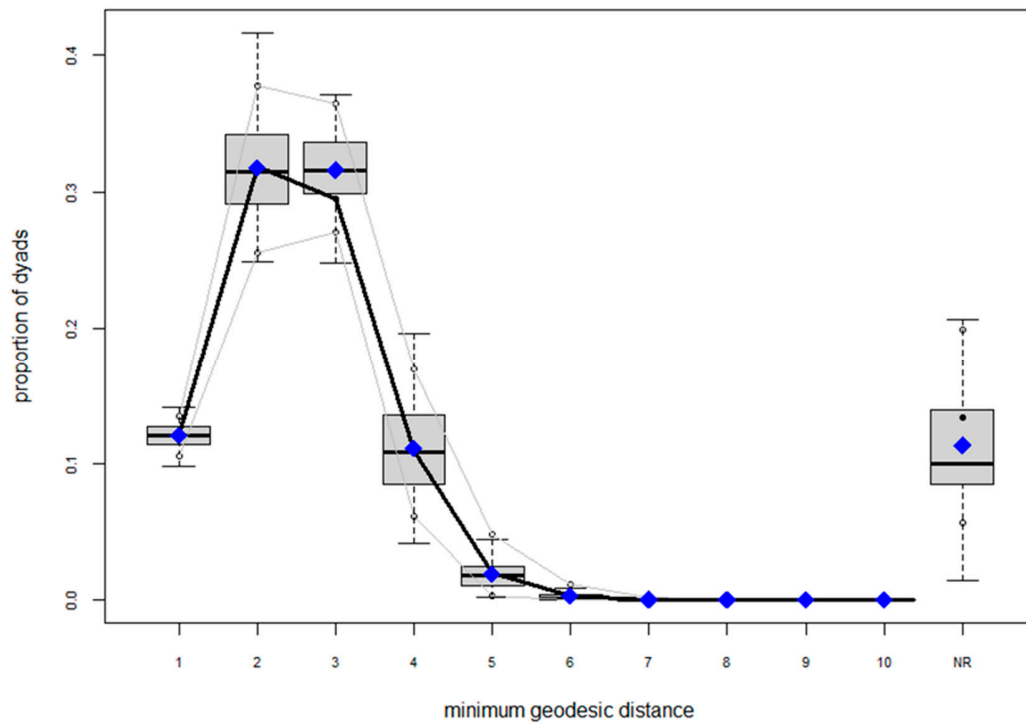


Figure S4. Goodness of fit plot for minimum geodesic distribution, preliminary model.

Goodness of Fit Plots
Final Model (Figures S5 – S8)

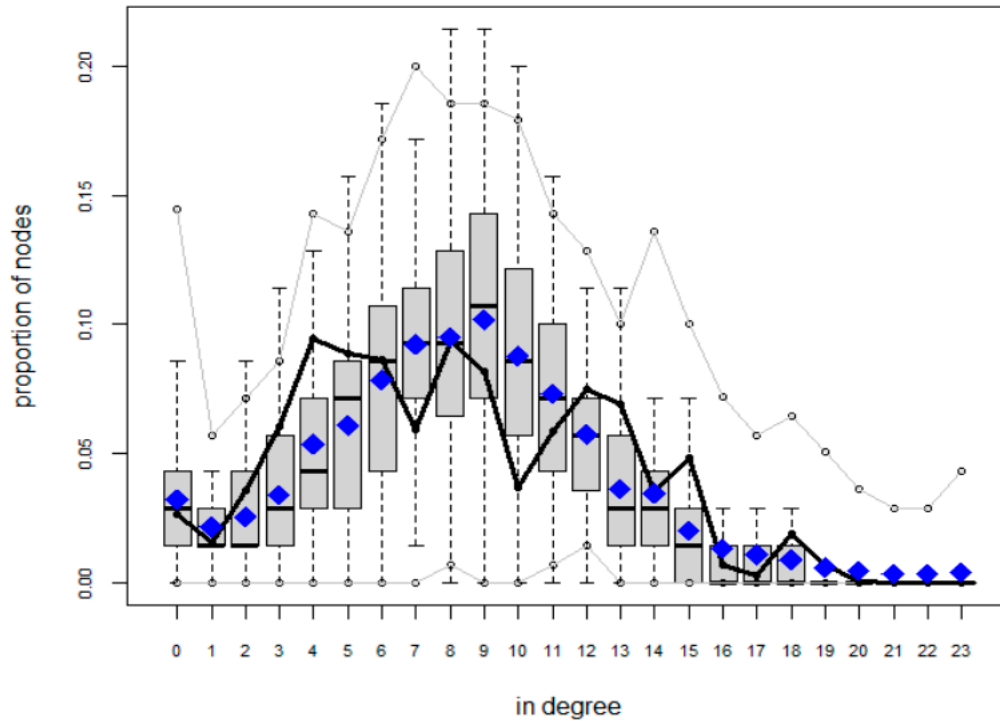


Figure S5. Goodness of fit plot for indegree distribution, final model.

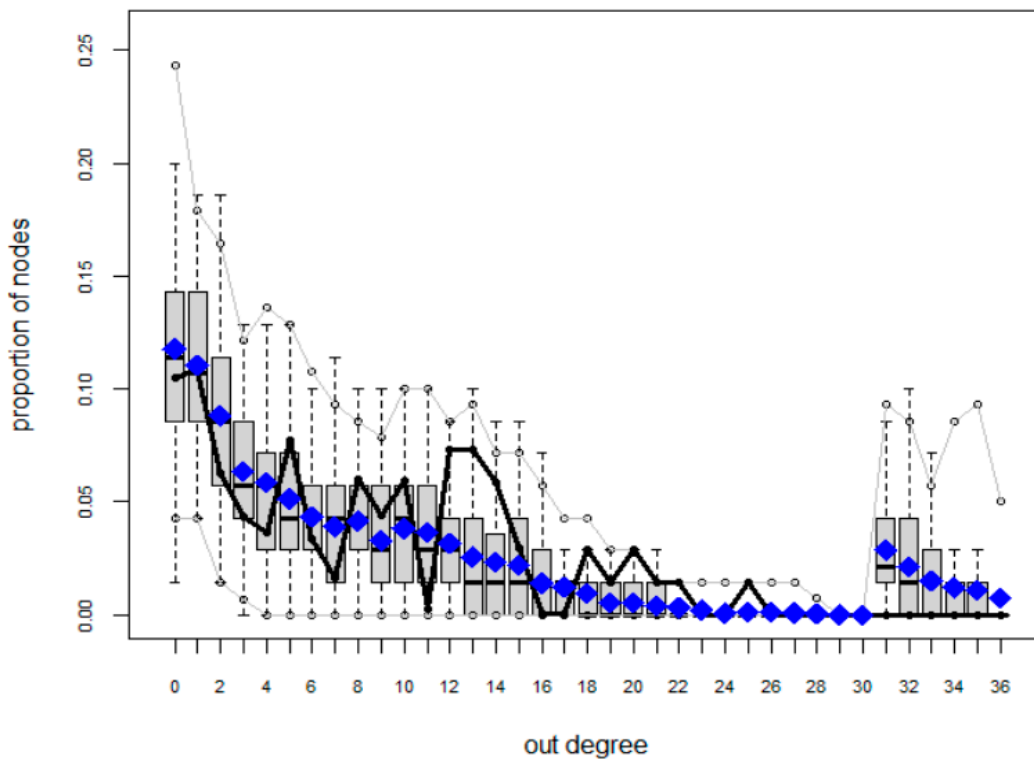


Figure S6. Goodness of fit plot for outdegree distribution, final model.

Goodness of Fit Plots
Final Model (continued)

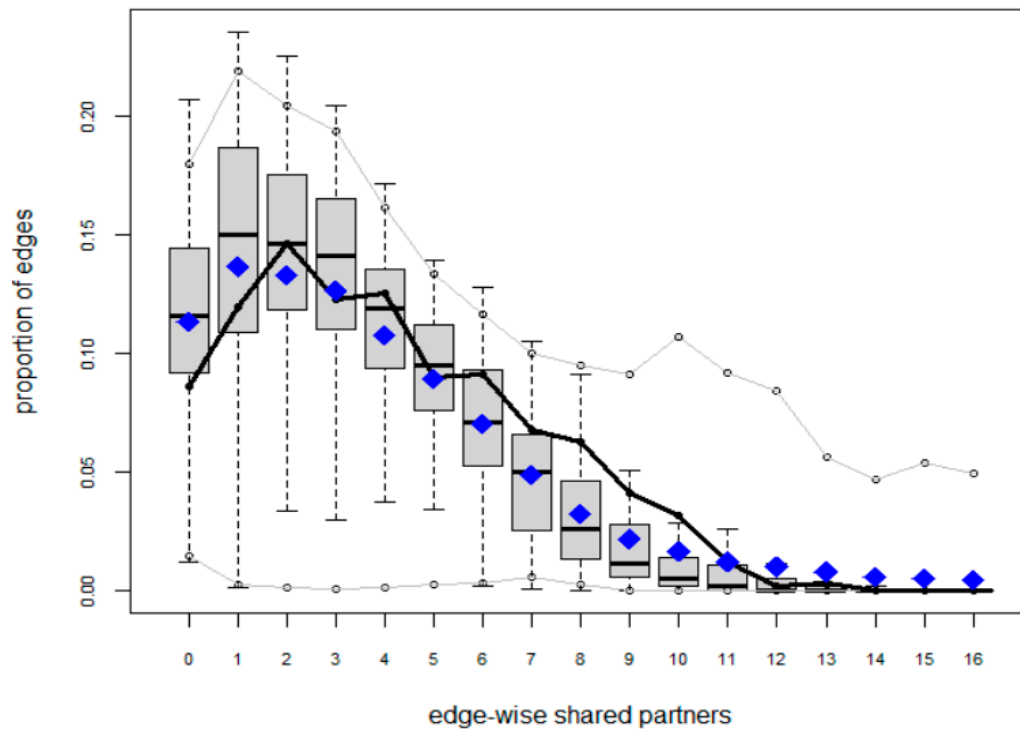


Figure S7. Goodness of fit plot for edge-wise shared partners distribution, final model.

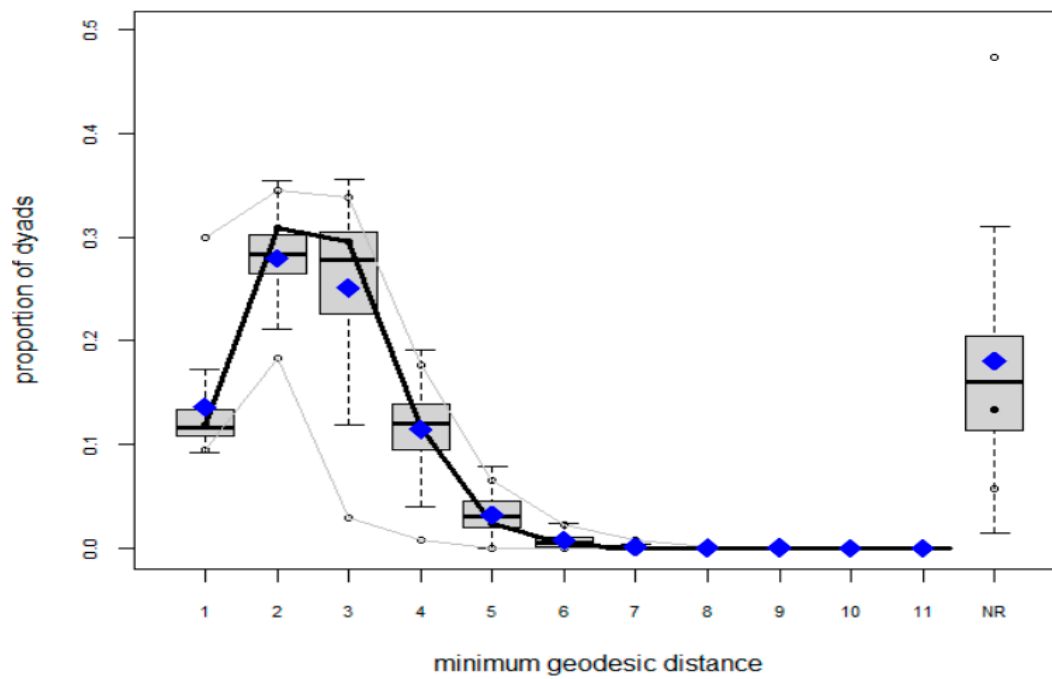


Figure S8. Goodness of fit plot for minimum geodesic distance distribution, final model.

Table S1. Preliminary curved exponential family (CEF) model results of factors associated with friendship ties among full-time Master of Social Work student ($N=70$) midway through the second year of a 2-year program.

Effect	Preliminary Model		
	θ	S.E.	p
Network structural effects			
Edges	-5.43	1.49	<0.001 ***
Reciprocity	2.32	0.19	<0.001 ***
Popularity spread (<i>gwidegree</i>)	-0.76	1.03	0.466
(decay parameter)	-0.31	0.80	0.693
Activity spread (<i>gwodegree</i>)	-1.36	0.37	<0.001 ***
(decay parameter)	1.38	0.40	<0.001 ***
Transitive shared partners (<i>gwesp</i>)	0.58	0.17	0.001 ***
(decay parameter)	0.19	0.18	0.287
Two-paths (<i>gwensp</i>)	-0.12	0.02	<0.001 ***
(decay parameter)	0.64	0.29	0.027 *
Individual-level factors			
Age			
outgoing ties	0.01	0.01	0.218
incoming ties	0.01	0.01	0.203
homophily ^a	-0.02	0.01	0.002 **
Gender (ref=male)			
female - outgoing ties	-0.51	0.22	0.019 *
female - incoming ties	-0.81	0.22	<0.001 ***
homophily ^b	0.78	0.19	<0.001 ***
Race/ethnicity (ref=White)			
Black/African American - outgoing ties	-0.17	0.12	0.154
Black/African American - incoming ties	-0.13	0.13	0.320
Hispanic/Latino(a) - outgoing ties	-0.24	0.23	0.279
Hispanic/Latino(a) - incoming ties	0.48	0.14	<0.001 ***
Other - outgoing ties	0.14	0.14	0.307
Other - incoming ties	0.43	0.22	0.047 *
race/ethnicity homophily ^b	0.41	0.10	<0.001 ***
Grade point average (GPA)			
outgoing ties	0.37	0.24	0.131
incoming ties	0.21	0.31	0.499
homophily ^a	-0.61	0.29	0.032 *
Dyadic and institutional factors			
Same cohort in first semester	0.86	0.12	<0.001 ***
Shared classes after first semester (0 - 11)	0.75	0.14	0.000
Proximity of home address proximity (within 10 km)	-0.01	0.13	0.938
AIC	2114		
BIC	2298		

Note. ^aHomophily is operationalized as absolute difference in value and is indicated by a significant negative parameter estimate; ^bHomophily is operationalized as same value of categorical variable and is indicated by a significant positive parameter estimate; *gwidegree* = geometrically weighted indegree; *gwodegree* = geometrically weighted outdegree; *gwesp* = geometrically weighted edgewise-shared partners; *gwensp* = geometrically weighted non-edgewise-shared partners; AIC = Akaike information criterion; BIC = Bayesian information criterion. * $p<0.05$; ** $p<0.01$; *** $p<0.001$.