Appendix C: Model goodness of fit

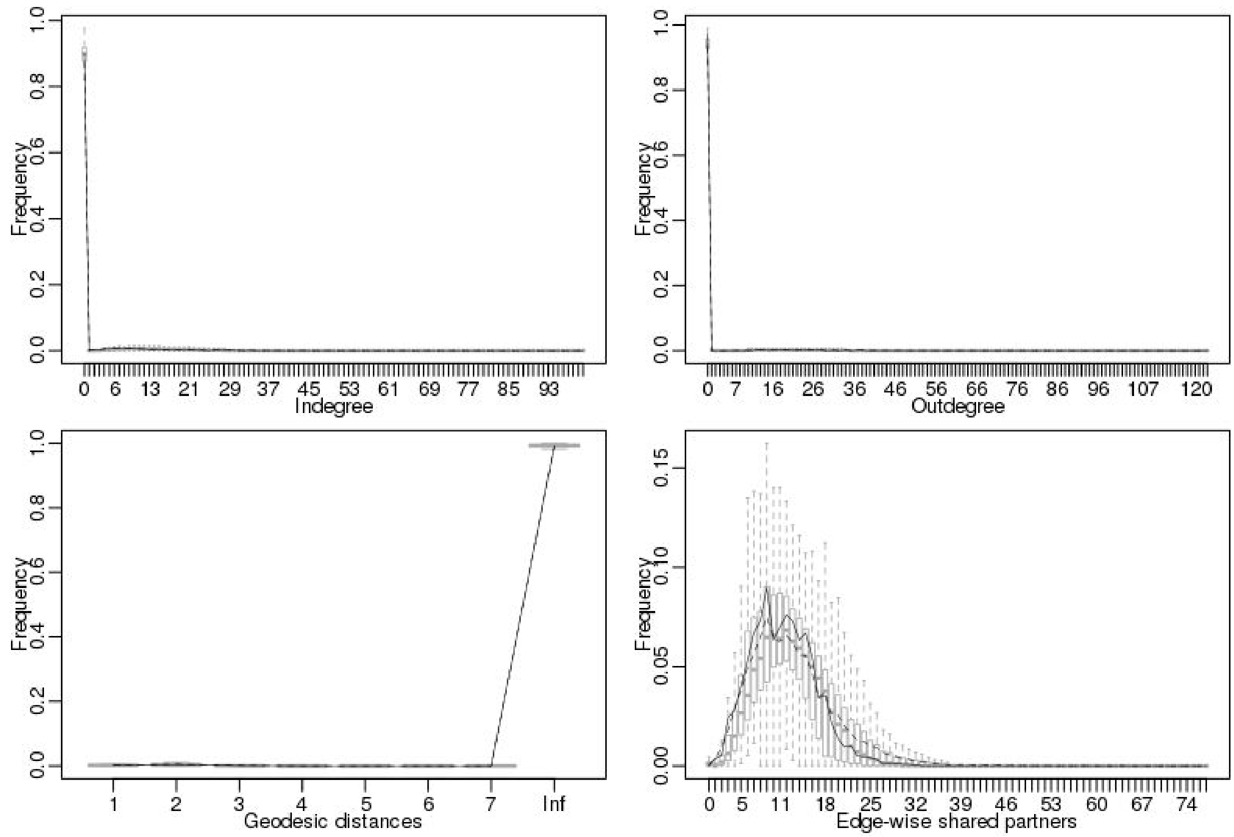

Figure C1: Comparison of observed and simulated network statistics generated by model 1.

As described in the methodology section of the manuscript, a TERGM facilitates inference by comparing the observed network to a distribution of simulated networks. Thus, it is important to confirm that the model is able to generate simulated networks that bear structural resemblance to the observed network (Koskinen and Snijders 2013). Figure C1 below compares four observed network statistics (outdegree, indegree, minimum geodesic distance [the minimum number of ties linking each pair of network actors], and edgewise shared partners) to a distribution of these same statistics based upon simulated networks generated by model 1. Overall, there is little discrepancy between the observed network (averaged across time steps) and simulated graphs. Because the overall network is quite sparse, and there are many high degree actors, the top two panels in figure A4 likely appear somewhat strange. The basic issue is that in both the simulated and observed networks, there are very few actors who have an indegree values of 45 or an outdegree value of 60. Because both the observed networks (dark line) and simulated networks have very few such actors at any given degree value, the frequency line remains at or near zero. The more important implication, however, is that the simulated networks do not have unwanted clusters of high degree actors, and that the distribution of indegree and outdegree values by node track resemble observed values.

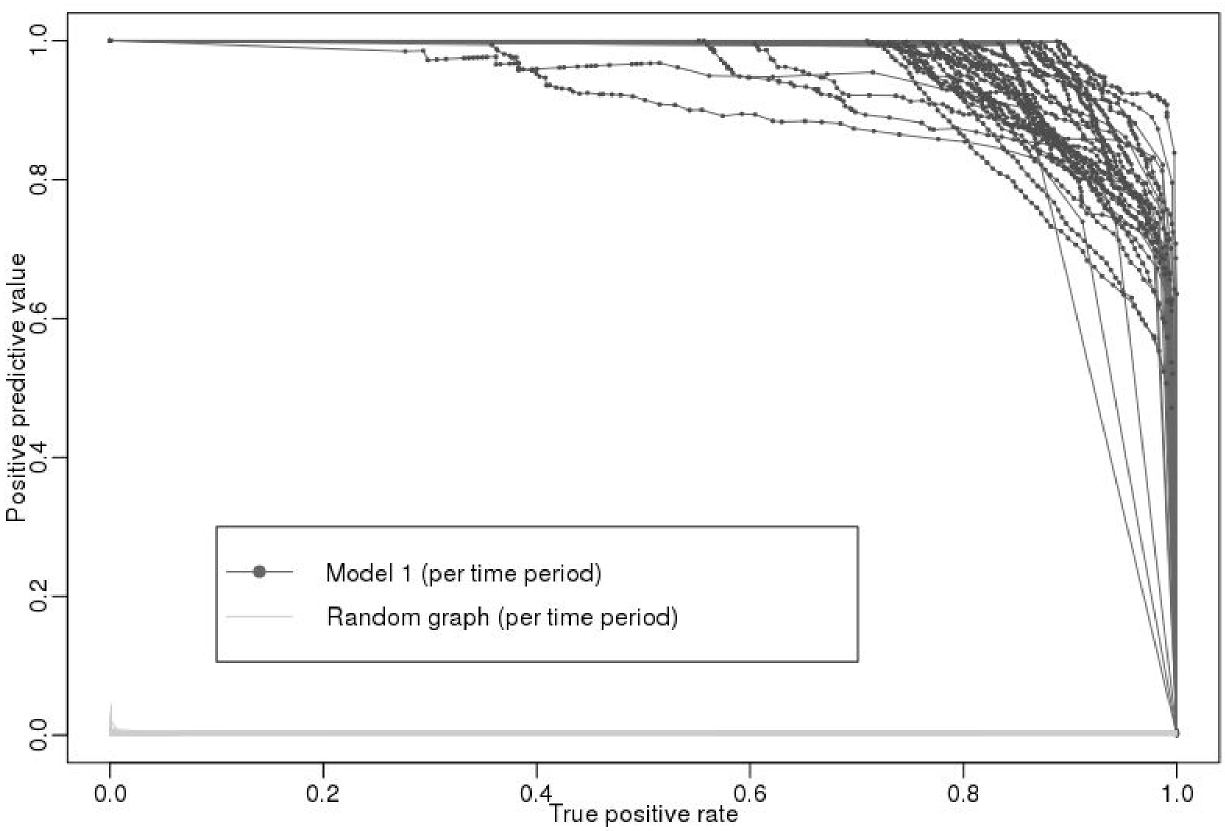

Figure C2: Precision-recall curves for model 1 and random graph by time period

A second way to gauge goodness-of-fit is to consider the predictive ability of the estimate model. Precision-recall curves compare the ability to correctly predict the presence of ties that are observed (the true positive rate is the proportion of observed ties correctly predicted by the model, i.e., “recall”) against the proportion of ties predicted by the model that are actually observed in the network (the number of true positives divided by the total number of predicted ties, i.e., “precision”). As a general rule, there is a tradeoff between precision (which is increased by making the model more discerning and reducing false negatives) and recall (which is increased by making the model less discerning and reducing false negatives), such that a precision-recall curve trends downwards. Figure C2 shows that model 1 performs very well in both regards, but still exhibits this tradeoff. Most importantly, compared to the precision-recall curves plotted for random graph estimates (i.e., an intercept-only model with just an edges to to control for density), model 1 performs far better.