Sample Size Recommendations for Estimating Cross-Sectional Network Models

Mihai A. Constantin

Tilburg University

Author Note

Coordinator: Angélique O. J. Cramer

Abstract

Network models are increasingly popular for describing psychological phenomena. Recent years translated into rapid methodological advancements, a significant amount of work being dedicated to assessing the accuracy of estimated network parameters, after the data has been collected. In the current work we suggest an approach, aimed at researchers who wish to know beforehand what sample size is roughly needed in order to accurately estimate a hypothesized network structure. We employ a simulation study with five design factors (i.e., sample size, network architecture, network connectedness, number of nodes, and type of data), and assess the estimation performance looking at three indicators: sensitivity, specificity, and edge weights correlation. Ours results show that, for both binary and ordinal data, the estimation methods work well and, although, not all edges are successfully retrieved (i.e., moderate sensitivity), those retrieved can be generally considered accurate (i.e., high specificity). This is further confirmed by a high edge weights correlation coefficient. We found that sample sizes ranging from 250 to 350 are generally enough to observe moderate sensitivity, high specificity, and high edge weights correlation, when the networks are sparse and consist of 20 nodes or less. The simulation design and steps discussed in this study are implemented as a freely available R package called $netPower$.

Keywords: network models, sample size, accuracy, cross-sectional data, Ising, GGM

Sample Size Recommendations for Estimating Cross-Sectional Network Models

The network approach to psychopathology serves as an increasingly popular framework aimed at explaining the onset and maintenance of mental disorders (Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011; Boschloo et al., 2015; Cramer, Waldorp, van der Maas, & Borsboom, 2010; Fried et al., 2017). Within this framework, one can investigate the extent to which a set of variables (i.e., nodes) are connected to one another (i.e., edges) in a network structure. In the field of psychopathology, networks are based on the premise that symptoms of a disorder act as a complex dynamical system that casually influence one another via direct relationships (Borsboom & Cramer, 2013; Cramer et al., 2016; Schmittmann et al., 2013).

Contrary to other fields where networks are constructed (e.g., in social networks an edge between two nodes may indicate a friendship relationship), networks describing psychological behavior (i.e., psychological networks) are typically estimated. In the case of psychological networks, nodes represent observed variables (e.g., symptoms of a disorder) and edges denote statistical relationships between these observed variables. Hence, in order to numerically encode the relationship between two nodes, this relationship must be estimated under specific statistical assumptions (Borsboom & Cramer, 2013; Epskamp & Fried, 2018). The result of this estimation is a so-called *network structure* or *graph* that can be visually inspected and analyzed by means of network inference indicators derived from graph theory (see Epskamp, 2017).

Being able to accurately estimate statistical parameters is considered to be of the utmost importance in network research of clinical phenomena, currently a subject of open debates and catalyst of rapid methodological advancements (Borsboom et al., 2017; Borsboom & Cramer, 2013; Epskamp, Borsboom, & Fried, 2018; Fried & Cramer, 2017). Accurate estimations are an

aspiring standard intended for safeguarding against erroneous conclusions and, as such, aims to contribute to the robustness and replicability of psychological research (Open Science Collaboration, 2015). Current network research embraces this view, declaring it a focal point in order for the young field of clinical network modelling to start ‘off the right foot’ (Epskamp et al., 2018, p. 196). In the present paper we subscribe to this view by exploring the question of accuracy as an interplay between sample size and estimated network parameters.

The remainder of this paper is structured as follows. First, we outline two popular network models for continuous and binary data and we discuss a commonly used regularization technique in order to avoid false positives (i.e., estimating relationships between nodes that in reality are not present). Next, we present recent work on the topic of reliable estimation of network parameters and discuss how the approach proposed in this paper complements it. Then, we proceed by presenting the design of the simulation study. Finally, we conclude by discussing limitations and points for improvement.

Pairwise Markov Random Field Models in Psychopathology

This paper is focused on pairwise Markov random field (PMRF) models for continuous and binary data. PMRF models are a popular subclass of network models in clinical research, readily applicable to a large number of cross-sectional datasets (Costantini et al., 2015; van Borkulo et al., 2014). PMRF models represent conditional (in)dependence relations: two nodes are only connected to one another if they are dependent conditional on all other nodes in the network. In other words, for two nodes to be connected that relationship cannot be explained by any other nodes in the network. In addition, PMRF models have two properties particularly useful in the context of psychological research. First property highlights that conditional dependence between two variables is consistent with a causal hypothesis, but not sufficient (see

Pearl, 2000). Second, PMRF models benefit from being well defined (i.e., a PMRF does not have any other equivalent models) making the interpretation less ambiguous (Epskamp, 2017).

Given that the data are continuous and follow a multivariate normal distribution, the appropriate PRMF model is the Gaussian Graphical Model (GGM; Costantini et al., 2015; Epskamp & Fried, 2018), also known as a *concentration network*, or a *partial correlation network*. In this case, conditional independence between two nodes is exactly equivalent to a partial correlation of zero. A connection (henceforth: edge) between two nodes reflects these partial correlations: the thicker the edge, the higher the partial correlation (henceforth: *edge weight*; see Epskamp & Fried, 2018). The edge weights can be obtained either from the inverse of a variance-covariance matrix, or through node-wide regressions, in which case each node is regressed on all other nodes (Epskamp & Fried, 2018; Højsgaard, Edwards, & Lauritzen, 2012).

When data are binary, the appropriate PMRF is the Ising model (van Borkulo et al., 2014), which holds a similar interpretation as the GGM. This type of PMRF model is based on a series of pairwise logistic regressions, resulting in two parameters being estimated for each node. The first parameter estimated is the interaction parameter denoted by β_{jk}, which encodes the edge weight between nodes j and k. The second parameter is the threshold parameter denoted by τ_j which indicates the activation tendency of node j, regardless of neighboring nodes. In the case of Ising models, drawing an edge between any two nodes is further influenced by the type of rule selected. When the AND-rule is used, an edge is drawn between nodes j and k only if both regression coefficients (i.e., interaction parameters) β_{jk} and β_{kj} of the regularized solution (i.e., to be explained in detail later) are estimated to be nonzero. In this case, the edge weight is given by the average of both parameters. When using the OR-rule, only one coefficient is required to

be nonzero for the edge to be drawn. Compared to OR-rule, the AND-rule is more conservative and results in a sparser network structure.

Regularization of Network Models

As with all statistical models that require the estimation of a sizeable number of parameters, sampling variation and limited sample sizes are significant threats to the accurate estimation of PMRF network models. As mentioned earlier, conditional independence between two nodes is equivalent to a partial correlation coefficient of 0. However, due to sampling variation, an estimated edge weight will most likely never be exactly 0, but very close around it. This results in multiple weak edges that are included in the network structure, that may actually prove to be *spurious*—edges that are not present in reality (Costantini et al., 2015; Epskamp, 2017). An attempt to overcome the issue of estimating spurious edges is to fit *regularized* PMRFs models. A popular regularization technique, demonstrated to work well for both continuous and binary data, is the *least absolute shrinkage and selection operator* (LASSO; Epskamp et al., 2018; Tibshirani, 1996; van Borkulo et al., 2014).

LASSO entails a two-step approach. In the first step, a range of networks are estimated with varying degree of sparsity given by a logarithmically spaced range of tuning parameters λ (lambda; see Epskamp & Fried, 2018; Zhao & Yu, 2006). During the second step, the network structure that provides the best fit to the data is selected from the range previously estimated. This is typically done by minimizing the Extended Bayesian Information Criterion (EBIC; Barber & Drton, 2015; Epskamp et al., 2018; Foygel & Drton, 2011). The EBIC is based on a hyperparameter γ (gamma) that controls the preference for a sparser network model. The value of the hyperparameter γ is the responsibility of researcher and it influences the type of network structures that are retrieved. For instance, values close to 0 ‘err on the side of discovery’

(Epskamp et al., 2018, p. 6), leading to higher sensitivity, whereas values above $.5$ indicate a preference for a more parsimonious model, thus resulting in higher specificity.

An alternative approach to overcome spurious edges is by only drawing edges based on significant partial correlation coefficients (Drton & Perlman, 2005). This implies conducting multiple significance tests while correcting the α level (e.g., Bonferroni correction) in order to keep the Type I error rate under control. However, given the rapidly increasing number of network parameters that need to be estimated, multiple testing needs to be conducted at extremely low α levels, increasing the Type II error and resulting in low statistical power (Epskamp et al., 2018). Therefore, LASSO is the current best practice in the network literature on psychopathology, given that the data are already collected.

Accurate Network Estimations: Before and After the Fact

The question of accuracy in the context of psychological networks arises from an interplay between the limited sample sizes often encountered in cross-sectional data and the network methodology with many parameters. While this problem naturally occurs also when fitting other models, such as factor models, the main difference is that, generally, network models require the estimation of more parameters. For instance, an Ising model (van Borkulo et al., 2014) with 14 nodes results in estimating 91 edge weights parameters (i.e., $\frac{k\times(k-1)}{2}$, where k represents the number of nodes in the network) and 14 threshold parameters (i.e., one for each node in the network). This amounts to 105 parameters for just 14 variables, making it already painfully clear that one obviously cannot obtain reliable parameter estimates with only 30 participants.

In light of these concerns, Epskamp et al. (2018) proposed an extra step before drawing conclusions on estimated network parameters. Their proposal involves a step-by-step procedure

for assessing the accuracy and stability of estimated edge weights and centrality indicators. This methodology describes three levels of assessment, readily available in the *bootnet R* package (Epskamp et al., 2018). This methodology is based on bootstrapping techniques, representing an approach that can be applied, given that the data are already collected. For example, during the first level of assessment, one can investigate how accurately the edge weights were estimated by constructing a 95% CI interval around each estimated edge-weight, such that, in 95% of the cases, the true value of an edge weight will fall in that interval.

The methodology introduced by Epskamp et al. (2018) stands out as an important tool for gaining insight into the accuracy of estimated network parameters. This methodology can be generally categorized as an *after-the-fact* approach, as it is designed for researchers who want to investigate the accuracy of estimated network parameters *after* the data have been collected. Simulation studies show that network models fit on larger sample sizes lead to more accurate estimates (Epskamp et al., 2018; Epskamp & Fried, 2018; van Borkulo et al., 2014). But what about the increasingly common situation that researchers, with a network approach in mind, plan to collect new data and wish to know *beforehand* what sample size is roughly needed in order to accurately estimate a network structure? This is exactly the question this paper aims to answer within the context of the Ising and GGM PMRF models.

The Current Study

In order to determine the adequate sample size, one is required to have an *a priori* expectation of certain aspects of the network structure, for example the number of nodes and the expected sparsity of the global structure. In this paper, we aim at investigating, by means of a simulation study, what sample sizes are needed to accurately retrieve network structures generated across a broad range of conditions (e.g., number of nodes, sparsity, architecture etc.).

We suggest and implement a *before-the-fact* approach as a simulation design with several steps. We start by generating model parameters for various network structures created under specific conditions. For example, assuming binary data and a small world architecture with 10 nodes, this step results in 45 interaction and 10 threshold model parameters. These parameters represent the *true network parameters*, or, simply, the *true model*, indicating that the exact data generating parameters are known. In the next step, we use these parameters to simulate data of varying sample sizes, say 100, 200, and 300. Next, we use these simulated data to estimate state-of-the art Ising models, or GGM when data is assumed ordinal. Finally, for each sample size, we investigate the accuracy of these *estimated* parameters as compared to the *true* network parameters. Ultimately, we aim at providing researchers with concrete sample size recommendations for accurately estimating cross-sectional network structures under a wide range of conditions.

Methods

The approach further discussed in this section is based on a simulation design with five factors. For each combination of factors and their levels (i.e., simulation condition, for example, a small-world architecture with 10 nodes, a rewiring probability of .5, and 100 participants) we estimated either the Ising model or GGM and monitored the deviations of the estimated parameters from the true model parameters. The simulation approach comprises five steps. The first step pertains to creating the simulation design by selecting the factors and their levels. The remaining four steps describe how the procedure was applied to every simulation condition, starting with: generating the true model parameters, generating the data, estimating the appropriate PMRF model, and computing the outcome measures. Each step is explained in detail below.

Simulation Design

The following factors and their respective levels—resembling but not identical to the Ising model validation study (van Borkulo et al., 2014)—were used to set up the broad range of simulation conditions under which performance of the Ising and GGM methods were evaluated given a certain sample size:

1. *Sample size*. Given that our primary goal is to assess the accuracy of network estimation at different sample sizes, we selected a wide, but fine-grained range of sizes. The sample sizes were picked between 50 and 1000 with increments of 50. This resulted in 20 levels as $S_{size} = \{50, 100, 150, \ldots, 900, 950, 1000\}$.
2. *Network size*. Three levels were used: $N_{nodes} = \{10, 20, 30\}$.
3. *Network architecture*. The network architecture was selected to reflect three commonly encountered network structures (see Figure 1), namely random (Barabasi & Albert, 1999), small-world (Watts & Strogatz, 1998), and scale-free (Cohen, Havlin, & Ben-Avraham, 2004).
4. *Network connectedness*. The level of connectedness was selected to generate sparse networks, the underlying assumption when using LASSO regularization, a technique almost always used when estimating cross-sectional network structures (see Epskamp & Fried, 2018; Epskamp, Kruis, & Marsman, 2017). For random networks, whether or not an edge is included in the network structure is a function of probability and three levels were selected: $P_e = \{.1, .2, .3\}$. For small-world networks, the network structure is generated iteratively as a function of two parameters: neighborhood (i.e., the number of neighboring nodes any given node is initially connected to), and rewiring probability (P_r). A sparse network was ensured by fixing the neighborhood parameter to 2, followed by

varying the level of P_r to obtain different network structures: $P_r = \{.1, .5, 1\}$. Scale-free network structures are also generated iteratively and modeled by two parameters: number of edges to be added per step, and preferential attachment (P_a). The number of edges to be added per step was fixed to one in order to obtain a sparse graph and the levels of P_a were varied as: $P_a = \{1, 2, 3\}$. Note that for small world and scale-free networks, changing the connectedness level did not result in denser network structures, as the neighborhood and edges per set were fixed values. However, in the case of random networks a higher P_e lead to more edges being included in the network structure, thus a denser network.

5. *Type of data*. Two data types were used, binary and ordinal.

The resulting design was a quasi-factorial design with 1080 cells $= 20$ (sample size) $\times 3$ (network size) $\times 3$ (network type) $\times 3$ (network connectedness) $\times 2$ (data type). Empirically, this can be understood as 27 different kinds of network structures (i.e., size, type, and connectedness), examined at 20 different sample sizes, both on binary and ordinal data.

Procedure

The following steps describe the order in which the procedure was applied to each simulation condition in the factorial design.

True Model Generation.

During this step, the true network structure (i.e., true model) was generated as follows:

- using the $igraph$ (Csardi & Nepusz, 2006) package in R version 3.5.1 (R Core Team, 2017) an undirected and unweighted graph was generated according to the configuration of each simulation condition, that is, based on three factors: network size, network architecture, and network connectedness. For example, a possible

generated network could have a size of 10 nodes and be of type random, in which case the connectedness would be operationalized as P_{edge}. Let the adjacency matrix describing the undirected unweighted generated network structure be represented as $\boldsymbol{\xi}$.

- next, model parameters were generated differently with respect to the appropriate PMRF based on the type of data. For both the Ising and GGM models, the number of edge weights to sample was determined as $\frac{k\times(k-1)}{2}$, where k stands for the network size factor (i.e., the number of nodes in the network). In the case of Ising model, a vector of edge weights $\boldsymbol{\beta}$ was sampled from the standard normal distribution with a mean of 0 and a standard deviation of 1. Then, a vector of threshold parameters $\boldsymbol{\tau}$ was sampled from the standard normal distribution with both the mean and standard deviation tied to $\boldsymbol{\beta}$, and with the sign inverted, in order to 'prevent nodes with many connections to be continuously activated and consequently having no variance' (van Borkulo et al., 2014, p. 8). Finally, the true network structure $\boldsymbol{\theta}$ was obtained by symmetrically mapping $\boldsymbol{\beta}$ onto the unweighted graph $\boldsymbol{\xi}$.
- a similar process was also used for generating GGM model parameters, with several additions. First, the edge weights $\boldsymbol{\beta}$ were uniformly sampled between 0 and 1 and, then, symmetrically mapped onto the unweighted graph $\boldsymbol{\xi}$, obtaining a matrix $\boldsymbol{\omega}$. To obtain a positive definite concentration matrix (i.e., inverse of a covariance matrix, or precision matrix), the algorithm described by Yin and Li (2011; see section 4.1), also used elsewhere (Epskamp, Rhemtulla, & Borsboom, 2017), was applied on the $\boldsymbol{\omega}$ matrix. Finally, the network true structure $\boldsymbol{\theta}$ was obtained by standardizing the $\boldsymbol{\omega}$ and

inverting the sign, thus obtaining a matrix of partial correlations that served as edge weights.

For both the Ising and GGM models, the edge weights were selected such that they only portrayed positive relationships between the nodes in the network. This is assumed to resemble empirical network structures of psychopathological phenomena (e.g., Boschloo et al., 2015; Fried, Epskamp, Nesse, Tuerlinckx, & Borsboom, 2016; McNally et al., 2015; Schmittmann et al., 2013).

Data Generation.

In this step, either a continuous or a binary dataset was generated, based on the true model parameters determined during the previous step. The resulting data were regarded as a sample from the unknown population described by the true parameters $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$. The number of observations matched the level of the sample size design factor for each respective condition (i.e., from 50 to 1000). For generating ordinal data only the $\boldsymbol{\beta}$ edge weights parameters were used, and the data were sampled from the multivariate normal distribution, as implemented in the $ggmGenerator$ function of the $bootnet$ (Epskamp et al., 2018) package in R. In this case, the data were generated with five item response steps, that is, resembling a Likert scale with five categories. When binary data were generated, the threshold $\boldsymbol{\tau}$ parameters were also used, alongside the $\boldsymbol{\beta}$ parameters, and the data were obtained using the $IsingSampler$ package (Epskamp, 2015). The binary data generating algorithm was the Metropolis-Hastings algorithm (Hastings, 1970).

In special cases where at least one node (i.e., column) in the sampled data had too little variance (i.e., a response category with lower than two observations), 10 resampling attempts under the same true model parameters were performed. If data still displayed little variance after

a total of 11 sampling attempts, the invariant nodes were removed from the data and this action recorded.

Model Estimation.

During this step, the estimated network parameters were obtained by fitting the appropriate PMRF model to the previously generated data. As a result, the estimated network structure $\widehat{\boldsymbol{\theta}}$ was obtained, described by the estimated edge weights $\widehat{\boldsymbol{\beta}}$, and, additionally, the thresholds $\hat{\boldsymbol{\tau}}$, when the data were binary. Both Ising and GGM models were fit with LASSO regularization. Recommended default values were used for the tuning parameters λ, and the EBIC hyperparameter γ, as implemented in the $IsingFit$ and $bootnet\ R$ packages (Epskamp & Fried, 2018; van Borkulo et al., 2014).

Outcome Measures.

In the final step of this procedure, the outcome measures were computed based on the deviations of the estimated network structure $\widehat{\boldsymbol{\theta}}$ from the true network structure $\boldsymbol{\theta}$. The primary outcome measure can be considered a binary classification for whether an edge is correctly estimated or not to be non-zero. Therefore, for each cell in the design, retrieval rates were investigated with respect to the edges in the true network structure and those in the estimated network structure. This implies looking at the proportions of the true/ false positive and negative estimated edges. First, the sensitivity and specificity of the estimation were computed (Adams & Leveson, 2012; Altman & Bland, 1994; van Borkulo et al., 2014). Sensitivity is regarded as the true positive rate (TPR) computed as:

$$sensitivity = \frac{\#\ true\ positive}{\#\ true\ positive + \#\ false\ negative}. \quad (1)$$

It may also be regarded as the power of the procedure, computed as $1 - \beta$, where β stands for Type II error rate, given by the false negative rate (FNR):

$$FNR = \frac{\#\ false\ negative}{\#\ false\ negative + \#\ true\ positive}. \quad (2)$$

In the context of this simulation, sensitivity is interpreted as the proportion of edges in the true network structure that were correctly estimated to be non-zero.

Specificity refers to the true negative rate (TNR) and provides information about the proportion of edges in the true network structure that were correctly estimated to be 0 (i.e., missing):

$$specificity = \frac{\#\ true\ negative}{\#\ true\ negative + \#\ false\ positive}. \quad (3)$$

Type I error rate was also computed, but not reported here, as the false positive rate (FPR), equivalent to $1 - specificity$:

$$FPR = \frac{\#\ false\ positive}{\#\ false\ positive + \#\ true\ negative}. \quad (4)$$

Additionally, correlations between the edge weights of the estimated model and those of the true model were also computed, in order to get a better understanding of how well they resemble each other (Epskamp & Fried, 2018). The edge weights correlation was computed between the upper triangle of the true network structure $\boldsymbol{\theta}$ and the upper triangle of the estimated network structure $\widehat{\boldsymbol{\theta}}$. In cases where the upper triangles resulted in unequal number of parameters (i.e., due to dropping nodes as a result of unsuccessful resampling attempts) the correlation coefficient was not computed, and replaced with a NA constant.

The procedure described above was applied to every cell in the factorial design and was replicated 100 times. That is, for each of the 1080 resulting conditions, we generated a true model, generated data based on the true model, estimated the appropriate PMRF on the data, and compared the estimated parameters to the true model parameters in terms of retrieval rates and edge weights correlations. This process was repeated 100 times for each condition, resulting in

estimating 54000 Ising models and 54000 GGM models. Figure 2 illustrates a true network structure and the corresponding estimated network structure for a specific simulation condition. The code used in this paper is freely available as an R package at https://github.com/mihaiconstantin/netPower. The installation, the data used in this paper, and the entire replication of the results can be achieved following the instructions provided in the README.md file.

Results

For each simulation condition, three outcomes were computed as a measure of performance: sensitivity, specificity, and edge weights correlation. In this section we report the means and standard deviations of each of these outcomes, after 100 replications. We start by outlining aggregated trends applicable to both binary and ordinal data, and then zoom in and present specific results related to each type of data. Figures 3 and 4 show the change in the mean of each outcome measure as a function of sample size, and with respect to the type of data, number of nodes, and network architecture. Figures 5 to 9 show the trends for each outcome measure using Tuckey boxplots for a better visual understanding of the spread in the outcome means (McGill, Tukey, & Larsen, 1978).

Sensitivity showed moderate mean values when averaged across all simulation conditions ($M = .615,\ SD = .280$) and increased steadily for each 50 participants added to the sample ($M = .035$), with larger gains for sample sizes smaller than 250 ($M = .093$), compared to larger samples ($M = .014$). Sensitivity did show more variation in the means for ordinal data ($SD = .309$) compared to binary data ($SD = .169$), result also clearly visible in the size of the boxplots under Figure 6. Furthermore, sensitivity was lower in the simulation conditions consisting of random architectures and 30 nodes ($M = .385,\ SD = .272$) compared to other

network architectures with the same number of nodes (scale free: $M = .665,\ SD = .289$; small world: $M = .531,\ SD = .250$; see Figure 5).

Specificity was very high when averaged across all conditions ($M = .942,\ SD = .077$). However, the largest values were observed for binary data ($M = .990,\ SD = .012$), where across all sample sizes the values generally remained similar (see Figure 8). Contrary, the GGM estimation produced somewhat lower values ($M = .895,\ SD = .085$). Additionally, GGM also resulted in lower specificity for networks with fewer nodes, dropping below $.8$. Figure 7 indicates that, in the case of scale free and small world networks, specificity was less stable as the sample size increased, the mean values being spread across a wider interval, however, it became more stable as the number of nodes increased.

Figure 4 shows that the correlation between the edge weights of the true model and the estimated model was high and displayed a stable pattern, increasing with sample size across all network architectures, for both binary and ordinal data ($M = .859,\ SD = .175$). On average, a sample size of 250 participants, was associated with a high correlation of $M = .818$ ($SD = .146$) and further increasing sample sizes lead to higher correlation coefficients and smaller variation. Figure 9 indicates, furthermore, that the means of the edge weights correlation coefficients were larger at smaller sample sizes when the network structure contained fewer nodes. Conversely, the coefficients were lowest for random networks with 30 nodes (see Table 5). The values reported so far on all three outcome measures show that, although some general trends can be depicted, such as increasing sensitivity, edge weights correlation, and all-around high specificity, these trends potentially vary with respect to the type of data used, therefore, requiring a closer inspection.

Binary Data

A closer look for models estimated on binary data indicates that sensitivity only reaches a moderate mean, when averaged across all conditions ($M = .488,\ SD = .169$). For sample sizes smaller than 450 and more than 20 nodes, sensitivity was low, regardless of the network structure ($M = .319,\ SD = .139$). However, for the same number of nodes, samples larger than 450 showed an almost double average increase in sensitivity with a very small standard deviation ($M = .580,\ SD = .037$). As presented in Table 2, sensitivity increased with $M = .029$ for each 50 simulated participants added to the sample size. This trend was larger in magnitude for samples smaller than 450 ($M = .051$), compared to sample sizes above 450 ($M = .012$), suggesting that the sensitivity stabilizes after a given number of participants. Furthermore, this trend was also uniformly present across all network structure, regardless of the level of connectedness, with a single exception. Namely, sensitivity was slightly lower for random networks with $P_e = .3$ ($M = .444,\ SD = .171$), as this type of operationalization also resulted in denser networks. Conversely, sensitivity was slightly larger for networks with fewer nodes: 10 nodes ($M = .538,\ SD = .160$), 20 nodes ($M = .486,\ SD = .164$), and 30 nodes ($M = .439,\ SD = .168$). Finally, we observed the highest values of sensitivity for sparse small world networks in combination with 10 nodes ($M = .579,\ SD = .150$).

Specificity was high all-around and we did not observe decreasing trends, regardless of sample size, network architecture, and operationalization of network connectedness (see Table 4). Only extremely low variation was observed when networks had 10 nodes ($M = .985,\ SD = .014$; see Figure 8). With respect to the edge weights correlation coefficients, averaging across all simulation conditions, we obtained high values with $M = .861$ and $SD = .134$. Bolded values in the right-hand side of Table 5 reveal that the correlation coefficients were lower at small

sample sizes (i.e., smaller than 200), falling below .8 ($M = .653,\ SD = .070$). On the other hand, for samples larger than 200, the mean of the edge weights correlation coefficients was very high, reaching $M = .913$ and $SD = .051$. This was observed across all network architectures and levels of connectedness. Similar to sensitivity, random networks with highest P_e resulted in lower edge weights correlation coefficients, as shown in Table 5.

Overall, the Ising model estimation showed moderate sensitivity, very high all-around specificity, and high edge weights correlation. Our simulation results indicate that the trend for sensitivity and edge weights correlation is to increase steadily as the sample sizes goes up, reaching values around .5 and .9, respectively, for 450 participants or more, regardless of the number of nodes and network architecture. These results align closely with current simulation work on fitting network models to binary datasets (van Borkulo et al., 2014).

Ordinal Data

In contrast to binary data, the GGM estimation on ordinal data displayed higher sensitivity values ($M = .742,\ SD = .309$). Table 1 shows that regardless of network architecture and number of nodes, samples sizes larger than 200 generally resulted in an average sensitivity value above .5 ($M = .830,\ SD = .217$), a threshold typically seen as 'acceptable as that at least indicates the strongest edges are discovered' (Epskamp & Fried, 2018, p. 11). A closer inspection of the results shows that scale-free networks had the highest sensitivity and lowest variation ($M = .912,\ SD = .172$). However, this result is subject to caution, as simulation conditions including this particular network structure resulted in simulation errors (see Table 6). Out of 18000 simulation runs that involved scale-free structures, 1092 could not be estimated (i.e., roughly 6%). Curiously, the error frequency went up as the number of nodes increased. We

backtracked the error to the data generation step of our procedure, where, on some occasions, datasets could not be samples based on the model parameters provided.

For the other network structures, the increase in sensitivity was linear and, on average, for each 50 participants, it went up by $M = .036$ (i.e., for random networks) and $M = .045$ (i.e., for small world networks). Sensitivity was higher for networks with fewer nodes: 10 nodes ($M = .876,\ SD = .194$), 20 nodes ($M = .735,\ SD = .295$), and 30 nodes ($M = .615,\ SD = .359$). Furthermore, Figure 6 also shows multiple outliers indicating that, for networks consisting of more nodes, empty structures were more often selected as best fitting by the EBIC model selection of LASSO regularization. The reason for this lies in the fact that more nodes resulted in denser network structures, which in turn resulted in lower generated partial correlation coefficients, making them more susceptible to LASSO. This was particularly the case for random networks, where sensitivity further decreased to 0 as the level of P_e increased (i.e., bolded values in Table 1 indicate this trend).

In contrast to the Ising model estimation, the GGM resulted in somewhat lower specificity, and larger variation. Table 3 shows that specificity decreased as the sample size increased (e.g., 50 simulated participants [$M = .978,\ SD = .043$] to 1000 simulated participants [$M = .867,\ SD = .085$]). This trend was particularly noticeable for small world and scale-free networks with 10 nodes, in which case the specificity dropped below .8. This finding was consistent across all levels of connectedness.

The correlation coefficient between the edge weights of the true model and estimated model was high when averaged across all simulation conditions ($M = .856,\ SD = .208$). For the smallest sample size of 50, the mean of the coefficients was low with $M = .430$ and $SD = .243$, however, it steadily increased with sample size, reaching $M = .958$ and $SD = .071$ for

sample sizes of 1000. Additionally, the values were overall higher for networks with fewer nodes, regardless of sample size: 10 nodes ($M = .922,\ SD = .114$), 20 nodes ($M = .867,\ SD = .181$), and 30 nodes ($M = .780,\ SD = .272$).

Overall, for ordinal data, the sensitivity was moderate to high and it increased with the sample size, at a faster rate in the case of scale-free and small world networks with fewer nodes. Specificity started high, but decreased as sample size increased, reaching values below .8 for networks with 10 nodes. Finally, the edge weights correlation was relatively high with low variation for sample sizes larger than 300. These results are generally, in line with other simulations (Epskamp, 2016), however, they are not directly comparable, as different design factors were employed.

Discussion

In this study we investigated what sample sizes are roughly needed for accurately estimating networks structures. To accomplish this, we employed a simulation approach based on five design factors: sample size, network architecture, network connectedness, number of nodes, and type of data. To capture the performance of the estimation we used recommended indicators in the literature on psychological networks, such as, sensitivity, specificity, and edge weights correlation (Epskamp & Fried, 2018; van Borkulo et al., 2014).

Rough Sample Size Recommendations

Our primary goal was to provide general recommendations for researchers interested in knowing beforehand what sample size are roughly needed for a hypothesized network structure. On one hand, binary data were estimated with lower sensitivity, but very high specificity all-around. On the other hand, ordinal data were estimated with higher sensitivity, but lower specificity. Furthermore, even though sensitivity was higher for lower sample size, it showed

greater variation, compared to the Ising estimation method. Nevertheless, in both cases, increasing sample sizes resulted in higher edge weights correlations. Table 7 presents an overview of the sample sizes recommended for accurately estimating various network structures.

Ordinal Data.

With respect to the number of nodes and network architecture, one can estimate a GGM model with 10 nodes and a random architecture with a sample size of 150 and expect the sensitivity, specificity, and edge weights correlation to be around .9. With 200 participants or more, for the same architecture, but 20 nodes, one can expect the sensitivity values to be around .8 and the specificity and correlation around .9, given that the network is sparse (i.e., $P_e = .1$). In case the network is somewhat denser (i.e., $P_e = .2$, or $P_e = .3$), similar values can be obtained with sample sizes larger than 600. When 30 nodes are used, sparse random networks (i.e., $P_e = .1$) can be estimated with around 500 participants, and still observe sensitivity values around .8, and specificity and edge weights correlation values around .9. For densely connected random networks with 30 nodes, one can either use more than 1000 participants in order to obtain moderate sensitivity and edge weight correlation values, or drop the number of nodes included, as suggested elsewhere (Epskamp et al., 2018). Nevertheless, even with a large number of nodes and a denser random network, one can expect the specificity to be close to 1.

When the researcher hypothesizes a small world network and wishes to study 10 variables, a sample size of 250 results in the best trade-off between sensitivity and specificity, both indicators reaching values close to .8, and a correlation larger than .8. Larger sample sizes in this scenario are associated with higher sensitivity, but fairly low specificity. For estimating small world networks with 20 nodes, at lest 500 participants are needed to ensure the sensitivity is around .8 and specificity and edge weights correlation around .9. For 30 nodes, similar values

can be expected with sample sizes larger than 650 participants. Furthermore, different levels of connectedness (i.e., $P_r = .1$, $P_r = .5$, and $P_r = 1$) did not appear to influence the values of the performance indicators.

Scale-free networks were associated with the lowest values for specificity and, conversely, highest values for sensitivity and edge weights correlation, along all networks structures. For a scale-free network with 10 nodes, one can obtain an optimal value around .8 for specificity and sensitivity (i.e., at the intersection of the two indicators) with only 100 participants, while still observing a correlation between the edge weights around .9. For 20 nodes, values around .9 on all outcomes indicators can be expected for samples larger than 150. For larger networks consisting of 30 nodes, similar values can be obtained with samples above 300 participants or more. Similar to small world networks, different levels of connectedness (i.e., $P_a = 1$, $P_a = 2$, and $P_a = 3$) did not result in different values for three outcomes measures investigated.

Binary Data.

Researchers fitting Ising models can expect stable and accurate estimations with respect to specificity. Regardless of network architecture, connectedness, number of nodes, and sample size, specificity is expected be very close to 1. When estimating a network structure with 10 nodes, one needs at least 350 participants for obtaining moderate sensitivity values of .5, but can expect a high edge weights correlation value around .9. For 10 nodes and binary data, the type of network architecture is not expected to influence the values of the outcome measures. For larger networks with 20 nodes, irrespective of their architecture, one can expect the sensitivity around .5 and the correlation close to .9, provided that data from more than 450 participants are

collected. When small world and scale-free networks with 30 nodes are estimated, one can expect moderate sensitivity and high correlation for 550 participants or more.

When the expected structure is a random architecture with 30 nodes, researchers also need to consider the overall network density. For instance, for an operationalization of $P_e = .1$, a sample size of 450 is adequate for an observed sensitivity of .5 and an edge weights correlation around .9. However, for denser networks with $P_e = .2$, in order to observe the same values, one needs to collected at least 550 participants. In the case of $P_e = .3$ we could not capture sensitivity values larger than .5, indicating the need for even larger sample sizes, or a different strategy with respect to the nodes included in the network structure.

Limitations and Future Directions

With this study we took a first step in answering multiple calls for tackling the issue of accurate network parameters estimations from the perspective of an a priori sample size analysis (Epskamp et al., 2018; Fried & Cramer, 2017). However, our approach reveals at least two limitations. A first limitation is suggested by the statement that 'the precise values of sensitivity, specificity and different correlations are strongly influenced by the expected network structure, similar to how the expected effect size influences a power analysis' (Epskamp & Fried, 2018, p. 11). More precisely, this reveals that, although we employed 1080 simulation conditions, we admittedly only scratched the surface, investigating just 27 different network structures. This limitation ought to encourage further studies to extend the range of network structures included in their simulation designs.

A second limitation concerns the way in which we operationalized the network connectedness factor. In line with previous studies we chose to investigate the idea of how edges are distributed in the network structure according to different algorithms (i.e., edge probability,

rewiring probability, and preferential attachment; van Borkulo et al., 2014). While we did not observe large changes in the means of our outcome measures for different levels of rewiring probability and preferential attachment, large changes were seen in the case of random networks, where edges were simply a function of probability. These changes were expected, as the probability of including an edge in the network is equivalent to the network density, thus, strongly influencing its sparsity.

For a first improvement, we suggest that further simulations also manipulate the network density across small world and scale-free architectures. This is particularly relevant for regularized PMRF models, where the assumption of sparsity is central to the LASSO methodology (Epskamp, Kruis, et al., 2017). On another note, during this study we constrained all edge weights to positive values, in order to resemble the generally positive relations between psychopathology symptoms. Further simulations may consider including an additional design factor for manipulating the positive ratio of the edge weights, adjusting the current approach to other fields, such as personality research (e.g., Costantini et al., 2015; Cramer et al., 2012). Additionally, $F1$ scores—which simultaneously take into account both precision and specificity (see; Jardine & van Rijsbergen, 1971; Powers, 2011)—can also be computed to account for situations in which sparse networks result in high specificity due to the ‘low base rate of connections’ (van Borkulo et al., 2014, p. 2).

Conclusion

In this paper we discussed a simulation approach suitable for investigating what sample sizes are roughly needed for arcuately estimating network parameters from cross-sectional data. We assessed the performance of 27 network structures at 20 sample sizes ranging from 50 to 1000, using state of the art network models for binary and ordinal data. Our simulation results

come close to those of other studies, showing that sensitivity and specificity increase to 1 as the sample size goes up (Epskamp, 2016; van Borkulo et al., 2014). This indicates that both Ising and GGM estimation methods work well and, although, not all edges are successfully retrieved (i.e., moderate sensitivity), those retrieved can be generally considered accurate (i.e., high specificity). This is further confirmed by the increasingly high edge weights correlation coefficient, suggesting that both methods also accurately retrieve the true edge values. As expected, both sensitivity and edge weights correlation values increased steadily with the sample size and were consistent across all three network architectures investigated. Generally, the GGM estimation required smaller sample sizes in order to correctly retrieve the edges and their weights, however, the retrieval rates were associated with higher variability. The Ising model, required larger samples to reach moderate sensitivity, however, the performance indicators showed only extremely low variability.

The approach discussed here is potentially useful for researchers with an idea about an expected network structure, looking for concrete sample size recommendations before the data collection plan is initiated. With respect to providing such concrete recommendations, we found that sample sizes ranging from 250 to 350 are generally enough to observe moderate sensitivity and high specificity and edge weights correlations, when the networks are sparse and consist of 20 nodes or less. Table 7 provides a complete overview of the sample size recommendations identified in this study. Furthermore, this approach can be complemented by established tools aimed at investigating the accuracy of estimated network parameters after the data is collected (i.e., bootnet; Epskamp et al., 2018). The simulation design and steps discussed in this paper are implemented as a freely available R package called $netPower$, which enables researchers to take informed guesses with respect to how many participants are needed for a hypothesized network

structure and a given number of nodes. Our tool, together with *bootnet*, is a first step in properly equipping applied researchers for getting started on the marathon of accurate parameter estimations in the increasingly popular field of psychological networks.

References

Adams, S. T., & Leveson, S. H. (2012). Clinical prediction rules. *BMJ (Clinical Research Ed.)*, *344*(11), d8312. https://doi.org/10.1136/bmj.d8312

Altman, D. G., & Bland, J. M. (1994). Diagnostic tests. 1: Sensitivity and specificity. *Bmj*, *308*(6943), 1552. https://doi.org/10.1136/bmj.308.6943.1552

Barabasi, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science (New York, N.Y.)*, *286*(5439), 509–512. https://doi.org/10.1126/SCIENCE.286.5439.509

Barber, R. F., & Drton, M. (2015). High-dimensional Ising model selection with Bayesian information criteria. *Electronic Journal of Statistics*, *9*(1), 567–607. https://doi.org/10.1214/15-EJS1012

Borsboom, D., & Cramer, A. O. J. (2013). Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The Small World of Psychopathology. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027407

Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., van Borkulo, C. D., van der Maas, H. L. J., & Cramer, A. O. J. (2017). False alarm? A comprehensive reanalysis of “Evidence that psychopathology symptom networks have limited replicability” by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, *126*(7), 989–999. https://doi.org/10.1037/abn0000306

Boschloo, L., Van Borkulo, C. D., Rhemtulla, M., Keyes, K. M., Borsboom, D., & Schoevers, R. A. (2015). The network structure of symptoms of the diagnostic and statistical manual of

mental disorders. *PLoS ONE*, *10*(9), 1–12. https://doi.org/10.1371/journal.pone.0137621

Cohen, R., Havlin, S., & Ben-Avraham, D. (2004). Structural properties of scale-free networks. *Handbook of Graphs and Networks*, (January), 85–110. https://doi.org/10.1002/3527602755.ch4

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. J. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13–29. https://doi.org/10.1016/J.JRP.2014.07.003

Cramer, A. O. J., van Borkulo, C. D., Giltay, E. J., van der Maas, H. L. J., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major Depression as a Complex Dynamic System. *Plos One*, *11*(12), e0167490. https://doi.org/10.1371/journal.pone.0167490

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., … Borsboom, D. (2012). Dimensions of Normal Personality as Networks in Search of Equilibrium: You Can't Like Parties if You Don't Like People. *European Journal of Personality*, *26*(4), 414–431. https://doi.org/10.1002/per.1866

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2–3), 137–150. https://doi.org/10.1017/S0140525X09991567

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Sy*, 1695. Retrieved from http://igraph.org

Drton, M., & Perlman, M. D. (2005). Multiple Testing and Error Control in Gaussian Graphical Model Selection. *Statistical Science*, *22*(3), 430–449. https://doi.org/10.1214/088342307000000113

Epskamp, S. (2015). IsingSampler: Sampling Methods and Distribution Functions for the Ising Model. *R Package Version 0.2.* Retrieved from https://cran.r-project.org/package=IsingSampler

Epskamp, S. (2016). Brief Report on Estimating Regularized Gaussian Networks from Continuous and Ordinal Data. Retrieved from http://arxiv.org/abs/1606.05771

Epskamp, S. (2017). *Network psychometrics*. University of Amsterdam, The Netherlands: Unpublished doctoral dissertation. Retrieved from https://dare.uva.nl/search?identifier=a76273c6-6abc-4cc7-a2e9-3b5f1ae3c29e

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*. https://doi.org/10.1037/met0000167

Epskamp, S., Kruis, J., & Marsman, M. (2017). Estimating psychopathological networks: Be careful what you wish for. *PLoS ONE*, *12*(6), 1–13. https://doi.org/10.1371/journal.pone.0179891

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized Network Psychometrics: Combining Network and Latent Variable Models. *Psychometrika*, *82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x

Foygel, R., & Drton, M. (2011). Bayesian model choice and information criteria in sparse generalized linear models. *Arxiv Preprint (ID 1112.5635)*. Retrieved from http://arxiv.org/abs/1112.5635

Fried, E. I., & Cramer, A. O. J. (2017). Moving Forward: Challenges and Directions for

Psychopathological Network Theory and Methodology. *Perspectives on Psychological Science*, *12*(6), 999–1020. https://doi.org/10.1177/1745691617705892

Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are "good" depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders*, *189*, 314–320. https://doi.org/10.1016/j.jad.2015.09.005

Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z

Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, *57*(1), 97. https://doi.org/10.2307/2334940

Højsgaard, S., Edwards, D., & Lauritzen, S. (2012). *Graphical Models with R*. New York: Springer-Verlag. https://doi.org/10.1007/978-1-4614-2299-0

Jardine, N., & van Rijsbergen, C. J. (1971). The use of hierarchic clustering in information retrieval. *Information Storage and Retrieval*, *7*(5), 217–240. https://doi.org/10.1016/0020-0271(71)90051-9

McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of Box Plots. *The American Statistician*, *32*(1), 12. https://doi.org/10.2307/2683468

McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental Disorders as Causal Systems. *Clinical Psychological Science*, *3*(6), 836–849. https://doi.org/10.1177/2167702614553230

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science.

Science, (6251), 349.

Pearl, J. (2000). *Causality Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Powers, D. M. W. (2011). Evaluation: From Precision, Recall and F-Measure To Roc, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63. https://doi.org/10.1.1.214.9232

R Core Team. (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*(1), 43–53. https://doi.org/10.1016/j.newideapsych.2011.02.007

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*. WileyRoyal Statistical Society. https://doi.org/10.2307/2346178

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, *4*, 1–10. https://doi.org/10.1038/srep05918

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Yin, J., & Li, H. (2011). A sparse conditional Gaussian graphical model for analysis of genetical genomics data. *The Annals of Applied Statistics*, *5*(4), 2630–2650. https://doi.org/10.1214/11-AOAS494

Zhao, P., & Yu, B. (2006). On Model Selection Consistency of Lasso. *Journal of Machine Learning Research*, *7*(Nov), 2541–2563. Retrieved from http://www.jmlr.org/papers/v7/zhao06a.html

Tables

Table 1.

Sensitivity as an outcome measure for ordinal data (i.e., GGM) and various simulation conditions: sample size (i.e., S_{size}), number of nodes (i.e., N_{nodes}), network architecture (i.e., random, scale-free, and small world), and connectedness (i.e., probability of an edge [P_e], preferential attachment [P_a], and probability of rewiring [P_r]).

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
50	10	.616	**.367**	**.140**	.543	.529	.563	**.030**	**.035**	**.034**
	20	**.052**	**.003**	**.001**	**.140**	**.065**	**.113**	**.004**	**.003**	**.001**
	30	**.002**	**.000**	**.000**	**.010**	**.004**	**.001**	**.000**	**.001**	**.001**
100	10	.923	.728	**.489**	.910	.885	.898	**.152**	**.158**	**.174**
	20	.544	**.052**	**.004**	.827	.821	.829	**.023**	**.018**	**.027**
	30	**.112**	**.002**	**.000**	.769	.781	.764	**.005**	**.016**	**.005**
150	10	.967	.882	.764	.924	.936	.947	**.403**	**.458**	**.417**
	20	.755	**.171**	**.014**	.878	.901	.884	**.093**	**.098**	**.087**
	30	**.344**	**.004**	**.001**	.865	.864	.865	**.024**	**.037**	**.043**
200	10	.981	.915	.815	.948	.946	.961	.654	.709	.658
	20	.822	**.378**	**.038**	.915	.921	.906	**.346**	**.292**	**.293**
	30	.540	**.023**	**.001**	.901	.901	.895	**.154**	**.136**	**.160**
250	10	.971	.943	.880	.949	.953	.962	.774	.768	.780
	20	.855	**.457**	**.093**	.927	.935	.931	.511	**.475**	**.478**
	30	.634	**.030**	**.001**	.915	.906	.912	**.382**	**.337**	**.293**
300	10	.973	.936	.896	.962	.967	.962	.820	.820	.819
	20	.886	.616	**.126**	.938	.937	.932	.654	.563	.591
	30	.693	**.067**	**.003**	.924	.929	.929	.574	**.485**	**.476**
350	10	.977	.954	.890	.970	.967	.979	.846	.838	.848
	20	.890	.696	**.249**	.947	.947	.945	.711	.660	.678
	30	.731	**.121**	**.004**	.932	.928	.926	.670	.560	.571
400	10	.990	.941	.926	.965	.965	.965	.844	.864	.871
	20	.910	.746	**.314**	.947	.953	.948	.765	.715	.718
	30	.770	**.168**	**.004**	.943	.943	.946	.707	.626	.621
450	10	.985	.958	.916	.976	.974	.967	.878	.875	.875
	20	.909	.768	**.484**	.958	.949	.953	.779	.754	.762
	30	.791	**.238**	**.009**	.950	.946	.946	.738	.683	.678
500	10	.983	.950	.933	.971	.975	.970	.878	.886	.880
	20	.921	.786	**.498**	.957	.954	.964	.796	.776	.777
	30	.810	**.335**	**.012**	.948	.941	.951	.766	.704	.699
550	10	.986	.964	.918	.966	.975	.972	.891	.899	.889
	20	.926	.804	.617	.959	.965	.958	.820	.800	.797
	30	.822	**.414**	**.009**	.949	.956	.951	.784	.732	.736

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
600	10	.986	.961	.930	.977	.972	.979	.892	.898	.902
	20	.925	.819	.693	.958	.961	.960	.842	.802	.807
	30	.832	**.482**	**.015**	.950	.951	.948	.789	.752	.733
650	10	.986	.959	.943	.978	.985	.982	.907	.904	.905
	20	.930	.834	.701	.956	.962	.962	.849	.821	.824
	30	.840	.527	**.018**	.956	.953	.955	.811	.769	.752
700	10	.993	.960	.934	.985	.978	.976	.903	.900	.900
	20	.942	.831	.735	.961	.962	.965	.855	.828	.833
	30	.861	.621	**.041**	.948	.957	.958	.819	.777	.770
750	10	.992	.968	.943	.978	.985	.974	.905	.917	.906
	20	.932	.859	.748	.970	.969	.966	.848	.842	.838
	30	.863	.602	**.081**	.957	.962	.959	.830	.788	.790
800	10	.990	.959	.945	.981	.982	.985	.914	.908	.927
	20	.940	.851	.754	.969	.966	.966	.867	.852	.843
	30	.874	.653	**.096**	.966	.968	.958	.837	.800	.802
850	10	.990	.971	.952	.986	.983	.982	.920	.919	.912
	20	.949	.865	.778	.966	.963	.969	.865	.844	.847
	30	.878	.677	**.112**	.966	.955	.959	.848	.807	.814
900	10	.986	.958	.950	.980	.980	.976	.921	.918	.925
	20	.944	.874	.783	.964	.970	.969	.877	.861	.866
	30	.873	.700	**.157**	.965	.959	.963	.846	.823	.826
950	10	.993	.979	.945	.974	.977	.983	.933	.920	.919
	20	.944	.874	.806	.971	.972	.978	.890	.865	.869
	30	.884	.713	**.243**	.966	.960	.967	.866	.827	.828
1000	10	.997	.976	.947	.975	.988	.976	.935	.929	.919
	20	.952	.881	.822	.971	.967	.973	.881	.877	.873
	30	.899	.716	**.306**	.962	.968	.961	.864	.834	.829

Note. Values of sensitivity smaller than .6 are in boldface.

Table 2.

Sensitivity as an outcome measure for binary data (i.e., Ising model) and various simulation conditions: sample size (i.e., S_{size}), number of nodes (i.e., N_{nodes}), network architecture (i.e., random, scale-free, and small world), and connectedness (i.e., probability of an edge [P_e], preferential attachment [P_a], and probability of rewiring [P_r]).

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
50	10	**.126**	**.130**	**.119**	**.134**	**.113**	**.136**	**.135**	**.143**	**.139**
	20	**.089**	**.085**	**.080**	**.092**	**.087**	**.079**	**.092**	**.081**	**.085**
	30	**.065**	**.056**	**.047**	**.063**	**.065**	**.069**	**.069**	**.059**	**.064**
100	10	**.285**	**.225**	**.252**	**.234**	**.206**	**.208**	**.251**	**.274**	**.267**
	20	**.183**	**.178**	**.169**	**.186**	**.185**	**.154**	**.174**	**.171**	**.181**
	30	**.156**	**.142**	**.095**	**.142**	**.139**	**.152**	**.164**	**.140**	**.144**
150	10	**.328**	**.302**	**.308**	**.284**	**.308**	**.304**	**.328**	**.356**	**.362**
	20	**.267**	**.263**	**.252**	**.244**	**.264**	**.249**	**.273**	**.252**	**.251**
	30	**.222**	**.203**	**.136**	**.216**	**.222**	**.226**	**.244**	**.222**	**.214**
200	10	**.406**	**.396**	**.407**	**.358**	**.341**	**.371**	**.421**	**.444**	**.444**
	20	**.320**	**.317**	**.335**	**.299**	**.318**	**.309**	**.352**	**.333**	**.343**
	30	**.283**	**.260**	**.198**	**.276**	**.289**	**.287**	**.312**	**.288**	**.304**
250	10	**.397**	**.412**	**.443**	**.415**	**.413**	**.409**	**.474**	**.499**	**.479**
	20	**.384**	**.379**	**.361**	**.356**	**.342**	**.365**	**.407**	**.392**	**.400**
	30	**.342**	**.324**	**.233**	**.345**	**.335**	**.335**	**.374**	**.334**	**.351**
300	10	**.447**	**.460**	**.492**	**.464**	**.459**	**.461**	.531	**.500**	.519
	20	**.406**	**.416**	**.414**	**.403**	**.383**	**.388**	**.449**	**.434**	**.451**
	30	**.383**	**.369**	**.220**	**.379**	**.369**	**.379**	**.408**	**.388**	**.380**
350	10	**.486**	**.497**	.537	**.468**	**.475**	.503	.573	.567	.569
	20	**.452**	**.450**	**.458**	**.434**	**.427**	**.422**	**.480**	**.471**	**.467**
	30	**.427**	**.400**	**.258**	**.410**	**.399**	**.395**	**.451**	**.432**	**.427**
400	10	.529	.515	.538	**.493**	.502	.522	.574	.585	.594
	20	**.486**	**.474**	.502	**.451**	**.445**	**.449**	.519	**.492**	**.495**
	30	**.456**	**.433**	**.282**	**.432**	**.449**	**.449**	**.490**	**.443**	**.462**
450	10	.557	.551	.576	.549	**.490**	.505	.591	.596	.618
	20	**.496**	.518	**.480**	**.493**	**.483**	**.486**	.547	.523	.527
	30	**.490**	**.457**	**.313**	**.464**	**.460**	**.451**	.504	**.479**	**.484**
500	10	.517	.581	.606	.512	.517	.559	.610	.644	.619
	20	.530	.536	.518	.514	.504	**.495**	.574	.542	.549
	30	.521	**.466**	**.333**	**.491**	**.479**	**.489**	.547	.506	.507
550	10	.525	.561	.615	.556	.570	.573	.643	.632	.639
	20	.554	.550	.557	.519	.523	.541	.593	.569	.566
	30	.521	.506	**.368**	.506	**.499**	.504	.570	.549	.540
600	10	.632	.599	.629	.589	.578	.578	.674	.680	.661
	20	.567	.559	.576	.527	.524	.548	.598	.589	.597

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
	30	.545	.514	**.332**	.539	.510	.518	.569	.553	.552
650	10	.598	.619	.641	.587	.593	.579	.670	.674	.685
	20	.577	.600	.603	.559	.559	.549	.615	.611	.607
	30	.557	.526	**.360**	.518	.529	.543	.602	.580	.569
700	10	.604	.608	.659	.623	.605	.590	.686	.685	.695
	20	.603	.609	.592	.558	.571	.558	.624	.615	.607
	30	.592	.554	**.330**	.551	.553	.554	.599	.592	.589
750	10	.616	.646	.655	.643	.623	.614	.700	.719	.702
	20	.598	.612	.611	.581	.598	.577	.640	.620	.630
	30	.593	.568	**.380**	.569	.552	.561	.624	.600	.594
800	10	.634	.687	.704	.644	.609	.626	.702	.724	.707
	20	.630	.658	.622	.580	.591	.593	.666	.641	.621
	30	.603	.572	**.363**	.574	.582	.571	.642	.619	.613
850	10	.644	.696	.700	.654	.636	.658	.727	.738	.710
	20	.627	.649	.629	.620	.603	.602	.677	.641	.649
	30	.602	.577	**.396**	.595	.589	.594	.649	.618	.622
900	10	.649	.648	.693	.649	.656	.647	.717	.729	.730
	20	.618	.646	.662	.608	.622	.628	.665	.668	.661
	30	.624	.590	**.394**	.587	.597	.591	.657	.637	.630
950	10	.653	.671	.694	.646	.676	.631	.740	.741	.730
	20	.640	.677	.663	.616	.625	.637	.687	.661	.682
	30	.640	.618	**.432**	.604	.616	.601	.672	.629	.632
1000	10	.638	.706	.723	.664	.638	.641	.736	.723	.746
	20	.666	.679	.669	.656	.630	.636	.695	.677	.672
	30	.647	.627	**.430**	.617	.617	.611	.670	.653	.661

Note. Values of sensitivity smaller than . 6 are in boldface.

Table 3.

Specificity as an outcome measure for ordinal data (i.e., GGM) and various simulation conditions: sample size (i.e., S_{size}), number of nodes (i.e., N_{nodes}), network architecture (i.e., random, scale-free, and small world), and connectedness (i.e., probability of an edge [P_e], preferential attachment [P_a], and probability of rewiring [P_r]).

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
50	10	.961	.958	.974	.863	.857	.876	.996	.994	.995
	20	1	1	1	.967	.996	.979	1	1	1
	30	1	1	1	.997	1	1	1	1	1
100	10	.920	.901	.925	**.792**	**.782**	**.782**	.982	.975	.961
	20	.981	.999	1	.936	.925	.910	1	1	1
	30	.998	1	1	.959	.953	.955	1	1	1
150	10	.931	.856	.859	**.780**	**.764**	**.785**	.915	.888	.911
	20	.968	.993	1	.914	.890	.900	.997	.996	.997
	30	.993	1	1	.937	.949	.950	1	1	1
200	10	.931	.867	.861	**.765**	**.776**	**.763**	.854	.833	.843
	20	.958	.978	.999	.899	.899	.909	.984	.987	.985
	30	.986	1	1	.940	.940	.935	.997	.997	.998
250	10	.916	.862	.814	**.771**	**.782**	**.765**	**.796**	.802	**.779**
	20	.953	.975	.995	.896	.885	.891	.973	.970	.972
	30	.980	1	1	.928	.940	.931	.991	.992	.994
300	10	.924	.873	.816	**.756**	**.748**	**.759**	**.791**	**.796**	**.773**
	20	.954	.962	.993	.878	.891	.899	.962	.963	.956
	30	.977	.999	1	.931	.921	.923	.982	.986	.987
350	10	.924	.868	**.794**	**.762**	**.755**	**.747**	**.766**	**.752**	**.770**
	20	.948	.950	.981	.888	.876	.880	.951	.950	.943
	30	.973	.997	1	.921	.925	.934	.979	.981	.980
400	10	.908	.856	.800	**.757**	**.765**	**.751**	**.769**	**.735**	**.760**
	20	.947	.941	.976	.877	.866	.880	.946	.942	.939
	30	.969	.995	1	.931	.922	.913	.975	.977	.979
450	10	.918	.843	**.799**	**.765**	**.746**	**.771**	**.757**	**.754**	**.739**
	20	.943	.932	.954	.871	.893	.862	.942	.930	.928
	30	.969	.994	1	.916	.920	.922	.971	.970	.971
500	10	.926	.851	**.787**	**.751**	**.741**	**.746**	**.737**	**.750**	**.752**
	20	.949	.931	.957	.889	.884	.859	.937	.926	.926
	30	.967	.989	1	.913	.932	.915	.968	.970	.969
550	10	.926	.856	**.784**	**.759**	**.739**	**.755**	**.744**	**.746**	**.753**
	20	.941	.929	.934	.867	.893	.877	.936	.924	.919
	30	.966	.985	1	.911	.920	.915	.966	.964	.963
600	10	.911	.836	.807	**.757**	**.760**	**.748**	**.734**	**.741**	**.741**
	20	.938	.922	.909	.886	.862	.884	.925	.917	.918

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
	30	.962	.977	1	.922	.911	.920	.965	.966	.966
650	10	.907	.843	**.781**	**.757**	**.746**	**.762**	**.751**	**.727**	**.732**
	20	.938	.910	.914	.877	.872	.868	.927	.912	.912
	30	.962	.976	1	.921	.925	.922	.963	.961	.963
700	10	.926	.854	**.790**	**.745**	**.732**	**.763**	**.748**	**.717**	**.741**
	20	.938	.914	.903	.886	.871	.875	.925	.904	.912
	30	.962	.966	1	.919	.913	.921	.962	.959	.960
750	10	.926	.836	**.788**	**.763**	**.747**	**.763**	**.742**	**.719**	**.735**
	20	.938	.911	.889	.877	.871	.865	.925	.911	.909
	30	.958	.969	.998	.913	.912	.913	.964	.960	.957
800	10	.917	.834	**.789**	**.748**	**.757**	**.754**	**.743**	**.720**	**.722**
	20	.937	.905	.895	.880	.877	.875	.919	.905	.904
	30	.960	.961	.997	.911	.907	.909	.962	.957	.956
850	10	.913	.831	**.784**	**.754**	**.736**	**.747**	**.730**	**.732**	**.703**
	20	.937	.906	.892	.861	.875	.863	.925	.910	.905
	30	.958	.958	.996	.914	.920	.913	.960	.957	.955
900	10	.902	.843	**.796**	**.725**	**.756**	**.760**	**.732**	**.726**	**.696**
	20	.939	.907	.885	.869	.875	.871	.915	.900	.899
	30	.961	.956	.994	.912	.911	.902	.960	.952	.953
950	10	.910	.843	**.780**	**.762**	**.764**	**.740**	**.738**	**.713**	**.712**
	20	.937	.905	.879	.852	.855	.865	.918	.899	.900
	30	.960	.953	.988	.907	.906	.907	.958	.953	.950
1000	10	.920	.839	**.772**	**.762**	**.733**	**.748**	**.713**	**.717**	**.697**
	20	.933	.902	.872	.863	.873	.860	.916	.899	.906
	30	.956	.952	.982	.910	.904	.915	.960	.952	.952

Note. Values of specificity smaller than .8 are in boldface.

Table 4.

Specificity as an outcome measure for binary data (i.e., Ising model) and various simulation conditions: sample size (i.e., S_{size}), number of nodes (i.e., N_{nodes}), network architecture (i.e., random, scale-free, and small world), and connectedness (i.e., probability of an edge [P_e], preferential attachment [P_a], and probability of rewiring [P_r]).

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
50	10	.995	.994	.993	.990	.994	.990	.988	.986	.991
	20	.999	.998	.995	.998	.998	.998	.998	.997	.998
	30	.999	.997	.985	1	1	1	.999	1	.999
100	10	.996	.996	.993	.995	.997	.993	.987	.984	.984
	20	.998	.997	.990	.998	.998	.999	.998	.997	.997
	30	.999	.996	.979	1	.999	1	.999	.999	.999
150	10	.996	.996	.990	.998	.995	.995	.983	.981	.979
	20	.999	.996	.986	.999	.998	.999	.996	.996	.996
	30	.999	.993	.974	1	1	1	.998	.998	.998
200	10	.997	.995	.987	.996	.996	.994	.976	.969	.979
	20	.999	.994	.984	.999	.999	.998	.996	.994	.994
	30	.999	.993	.967	1	1	1	.998	.998	.997
250	10	.997	.995	.987	.994	.994	.995	.974	.971	.971
	20	.998	.993	.980	.998	.999	.998	.995	.992	.994
	30	.998	.990	.961	.999	1	1	.998	.997	.997
300	10	.998	.993	.985	.995	.995	.995	.969	.972	.973
	20	.998	.993	.979	.998	.998	.999	.994	.991	.991
	30	.998	.988	.967	1	.999	.999	.997	.997	.997
350	10	.998	.993	.988	.997	.993	.994	.968	.964	.963
	20	.998	.993	.976	.998	.998	.999	.993	.992	.992
	30	.998	.989	.966	.999	1	.999	.997	.996	.996
400	10	.998	.995	.987	.995	.993	.992	.970	.959	.966
	20	.998	.993	.972	.999	.998	.998	.992	.992	.991
	30	.998	.987	.964	.999	1	.999	.997	.996	.995
450	10	.997	.993	.987	.995	.993	.994	.972	.966	.965
	20	.997	.991	.976	.998	.998	.997	.992	.991	.990
	30	.998	.986	.956	.999	.999	.998	.997	.996	.995
500	10	.997	.994	.983	.998	.993	.994	.963	.966	.963
	20	.998	.991	.976	.998	.998	.998	.991	.990	.989
	30	.998	.987	.955	1	.999	.999	.996	.995	.995
550	10	.998	.993	.984	.995	.994	.995	.962	.962	.973
	20	.998	.991	.973	.998	.998	.998	.991	.989	.990
	30	.998	.985	.949	.999	.999	.999	.996	.995	.995
600	10	.998	.994	.983	.995	.997	.994	.960	.957	.966
	20	.998	.989	.971	.998	.998	.998	.992	.989	.988

		Random			Scale-free			Small world		
S_{size}	N_{nodes}	P_e = .1	P_e = .2	P_e = .3	P_a = 1	P_a = 2	P_a = 3	P_r = .1	P_r = .5	P_r = 1
	30	.997	.984	.957	.999	.999	.999	.996	.994	.994
650	10	.998	.992	.988	.994	.996	.996	.963	.956	.957
	20	.997	.990	.966	.998	.998	.998	.991	.988	.989
	30	.997	.985	.958	.999	.998	.999	.995	.995	.994
700	10	.998	.993	.985	.993	.994	.994	.964	.963	.961
	20	.997	.988	.969	.998	.997	.997	.991	.988	.990
	30	.997	.984	.957	1	.998	.999	.996	.994	.994
750	10	.997	.992	.981	.995	.994	.995	.959	.953	.959
	20	.998	.989	.969	.998	.997	.998	.991	.988	.988
	30	.997	.983	.953	.999	.999	.999	.995	.994	.994
800	10	.997	.993	.984	.994	.994	.994	.963	.956	.961
	20	.998	.988	.965	.998	.998	.998	.990	.989	.989
	30	.997	.983	.954	.999	.999	.999	.995	.994	.993
850	10	.997	.992	.982	.995	.992	.995	.955	.961	.968
	20	.997	.990	.970	.997	.998	.997	.987	.988	.987
	30	.997	.983	.954	.998	.999	.999	.995	.994	.993
900	10	.997	.994	.983	.994	.996	.993	.965	.950	.959
	20	.998	.988	.963	.998	.997	.998	.991	.988	.987
	30	.998	.984	.952	.999	.999	.999	.996	.994	.994
950	10	.998	.992	.981	.995	.993	.993	.962	.960	.964
	20	.998	.988	.964	.999	.998	.998	.989	.987	.988
	30	.997	.979	.944	.999	.999	.998	.995	.994	.994
1000	10	.997	.992	.982	.994	.992	.994	.959	.959	.960
	20	.997	.989	.962	.997	.998	.998	.989	.986	.985
	30	.997	.979	.948	.999	.998	.998	.995	.993	.993

Table 5.

Correlation between edge weights as an outcome measure for ordinal (i.e., GGM) and binary (i.e., Ising model) data for various simulation conditions: sample size (i.e., S_{size}), number of nodes (i.e., 10, 20, and 30), and network architecture (i.e., random, scale-free, and small world).

	GGM									Ising								
	Random			Scale-free			Small world			Random			Scale-free			Small world		
Size	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30
50	.756	**.352**	**.129**	.745	**.612**	**.443**	**.331**	**.275**	**.225**	**.568**	**.443**	**.337**	**.550**	**.489**	**.433**	**.486**	**.446**	**.415**
100	.839	**.536**	**.261**	.873	.871	.850	**.549**	**.391**	**.309**	.719	**.622**	**.502**	.704	**.641**	**.607**	**.679**	**.616**	**.607**
150	.918	**.621**	**.400**	.927	.920	.916	**.695**	**.508**	**.426**	.788	.722	**.612**	.771	.748	.720	.758	.721	.707
200	.939	.701	**.476**	.945	.943	.940	.782	**.680**	**.574**	.843	.777	**.665**	.828	.799	.789	.824	.802	.781
250	.950	.727	**.478**	.959	.953	.953	.826	.763	**.694**	.859	.811	.716	.857	.840	.828	.857	.841	.820
300	.960	.776	**.521**	.965	.962	.958	.857	.826	.798	.894	.844	.734	.887	.863	.860	.875	.868	.850
350	.965	.806	**.546**	.969	.968	.966	.880	.865	.842	.904	.863	.772	.901	.881	.877	.897	.886	.874
400	.970	.845	**.561**	.974	.970	.971	.890	.890	.871	.914	.883	.784	.908	.895	.896	.905	.900	.892
450	.973	.879	**.602**	.976	.973	.974	.910	.904	.891	.925	.889	.801	.915	.915	.904	.918	.914	.905
500	.977	.886	**.640**	.978	.977	.976	.915	.915	.904	.929	.905	.821	.921	.922	.914	.926	.921	.913
550	.978	.908	**.655**	.981	.980	.977	.924	.924	.916	.937	.916	.832	.940	.930	.925	.932	.928	.924
600	.981	.923	**.672**	.984	.982	.980	.931	.931	.923	.944	.916	.826	.942	.934	.929	.943	.936	.927
650	.981	.926	**.686**	.983	.982	.983	.934	.939	.933	.946	.927	.841	.949	.942	.935	.944	.941	.937
700	.983	.932	.723	.985	.984	.983	.941	.943	.937	.950	.932	.833	.948	.944	.938	.951	.945	.940
750	.985	.939	.742	.988	.984	.983	.944	.947	.943	.956	.935	.854	.954	.947	.939	.951	.948	.942
800	.986	.943	.753	.988	.987	.983	.945	.951	.948	.962	.938	.857	.956	.950	.948	.954	.951	.948
850	.987	.949	.777	.988	.986	.986	.950	.953	.950	.960	.944	.858	.964	.955	.951	.959	.956	.950
900	.986	.951	.780	.989	.988	.986	.953	.957	.954	.962	.950	.862	.963	.959	.952	.961	.958	.956
950	.988	.955	.822	.990	.987	.986	.955	.960	.956	.966	.949	.881	.962	.961	.956	.965	.960	.956
1000	.989	.959	.833	.990	.987	.987	.958	.961	.959	.965	.955	.876	.965	.962	.959	.966	.963	.961

Note. Correlation are computed between the upper triangle of the true model (i.e., the model that generated the data) and the estimated model. Correlation coefficients smaller than .8 are in boldface.

Table 6.

Network structures that resulted in simulation errors and the frequency of occurrence.

N_{nodes}	Architecture	Connectedness	Data	Frequency
10	scale free	$P_a = 1$	ordinal (GGM)	24
10	scale free	$P_a = 2$	ordinal (GGM)	24
10	scale free	$P_a = 3$	ordinal (GGM)	20
20	scale free	$P_a = 1$	ordinal (GGM)	116
20	scale free	$P_a = 2$	ordinal (GGM)	115
20	scale free	$P_a = 3$	ordinal (GGM)	124
30	scale free	$P_a = 1$	ordinal (GGM)	199
30	scale free	$P_a = 2$	ordinal (GGM)	242
30	scale free	$P_a = 3$	ordinal (GGM)	228

Note. There were only 1092 errors out of 108000 simulation runs, at a rate of $.01$, occurring only for scale free networks, more frequently as the number of nodes increased.

Table 7.

Sample size recommendations for accurately estimating 27 network structure, from ordinal (i.e., GGM) and binary (i.e., Ising) cross-sectional data. In case of ordinal data, the sample sizes suggested are aimed at obtaining values of sensitivity and edge weights correlation around .8, and values of specificity larger than .9. For binary data, the sample sizes presented are aimed to obtain moderate sensitivity around .5, edge weights correlation around .8, and specificity larger than .9.

		GGM			Ising		
Architecture	Connectedness	10	20	30	10	20	30
Random	$P_e = .1$	150	200	500	350	450	450
	$P_e = .2$	200	550	900	350	450	550
	$P_e = .3$	200	600	**1000**	350	600	**1000**
Scale-free	$P_a = 1$	100	150	150	400	450	550
	$P_a = 2$	100	150	150	400	500	550
	$P_a = 3$	100	150	150	400	500	550
Small world	$P_r = .1$	250	500	650	300	400	450
	$P_r = .5$	300	550	800	300	450	500
	$P_r = 1$	250	600	800	300	400	500

Note. Values in boldface indicate that larger samples than those studied here are needed.

Figures

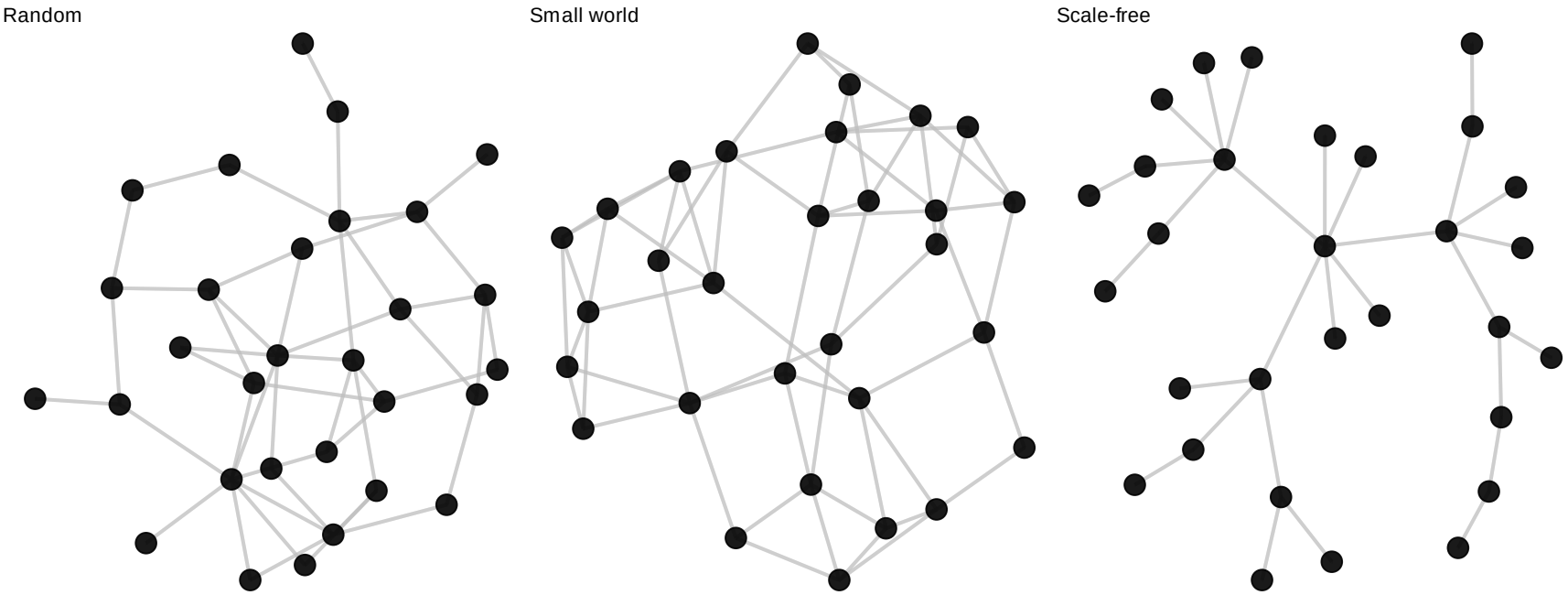

Figure 1. Example of unweighted graphs consisting of 30 nodes and representative for three different network architectures: random network (i.e., probability of an edge being included in the network structure is . 1), small world network (i.e., rewiring probability is . 1 and neighborhood parameter is 2), and scale-free network (i.e., power of prudential attachment and number of edges added per iteration are 1). Weighted variants of these networks were constructed by mapping generated model parameters onto them, obtaining true network structures against which the estimated network structures were compared.

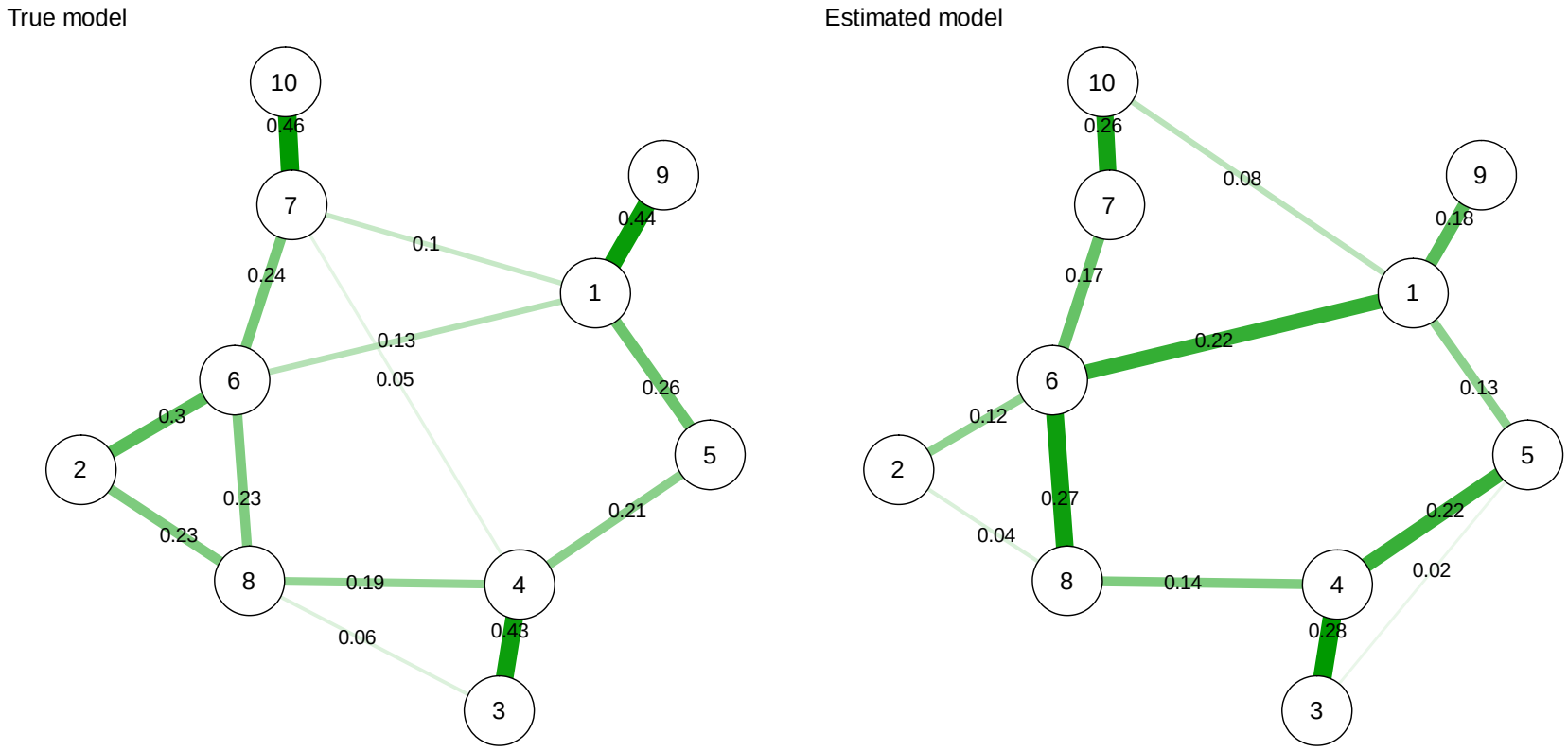

Figure 2. Example of a true network structure (i.e., on the left) and the estimated network structure (i.e., on the right). The true network structure corresponds to a random architecture, with 10 nodes and a probability of a connection of .2. The estimated network structure was obtained from simulated ordinal data consisting of 150 participants. The closer the resemblance between the two network structures, the more accurate the estimation. For this example, sensitivity is .785, specificity is .934, and the edge weights correlation is .873.

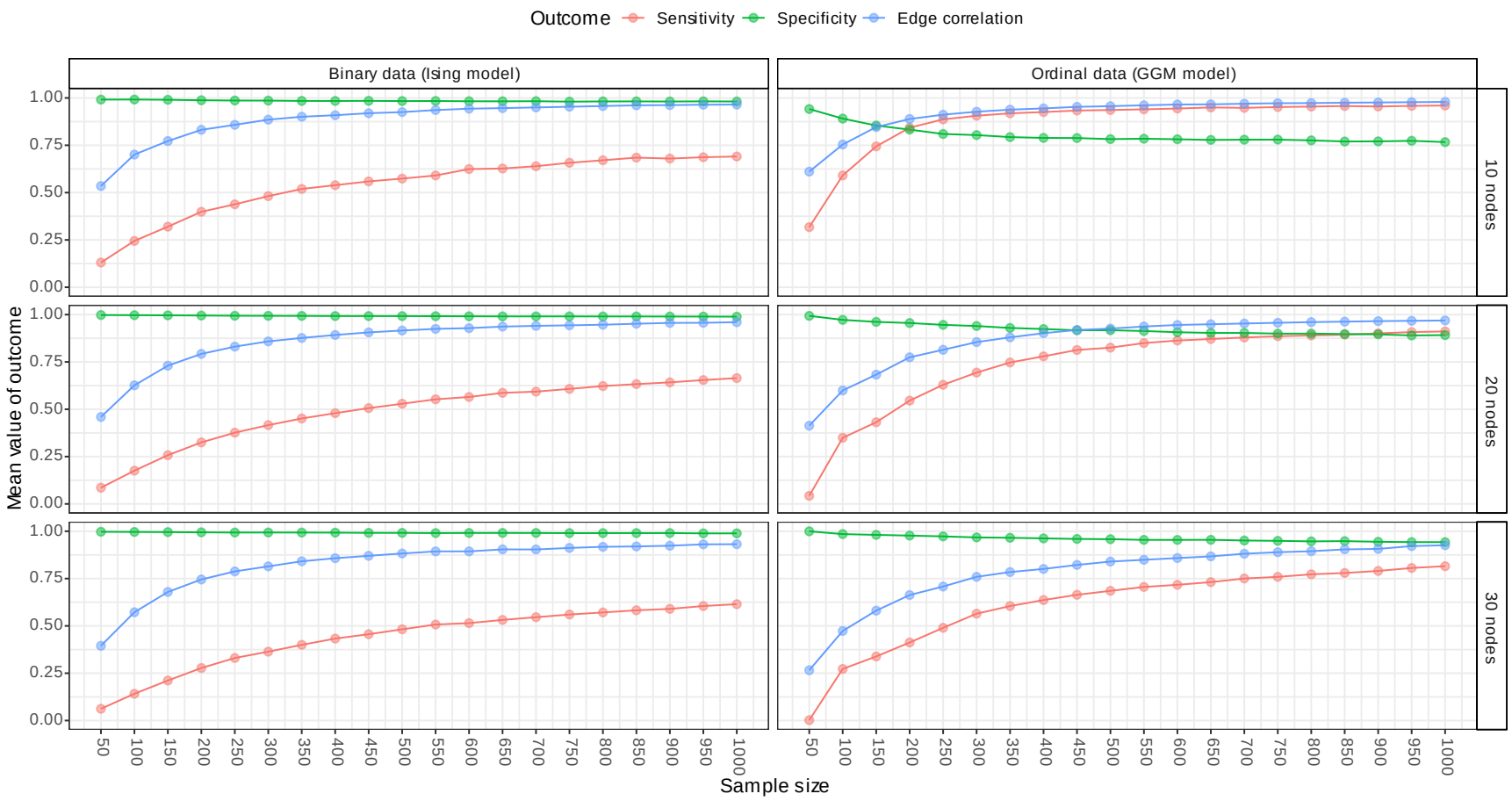

Figure 3. Sensitivity, specificity and edge weights correlation of simulated data. Horizontal panels indicate the type of data used and vertical panels the number of nodes in the network structure. The color of the lines indicates different outcome measures. The edge weights correlation was computed between the upper triangle of the data generating model (i.e., true model) and estimated model.

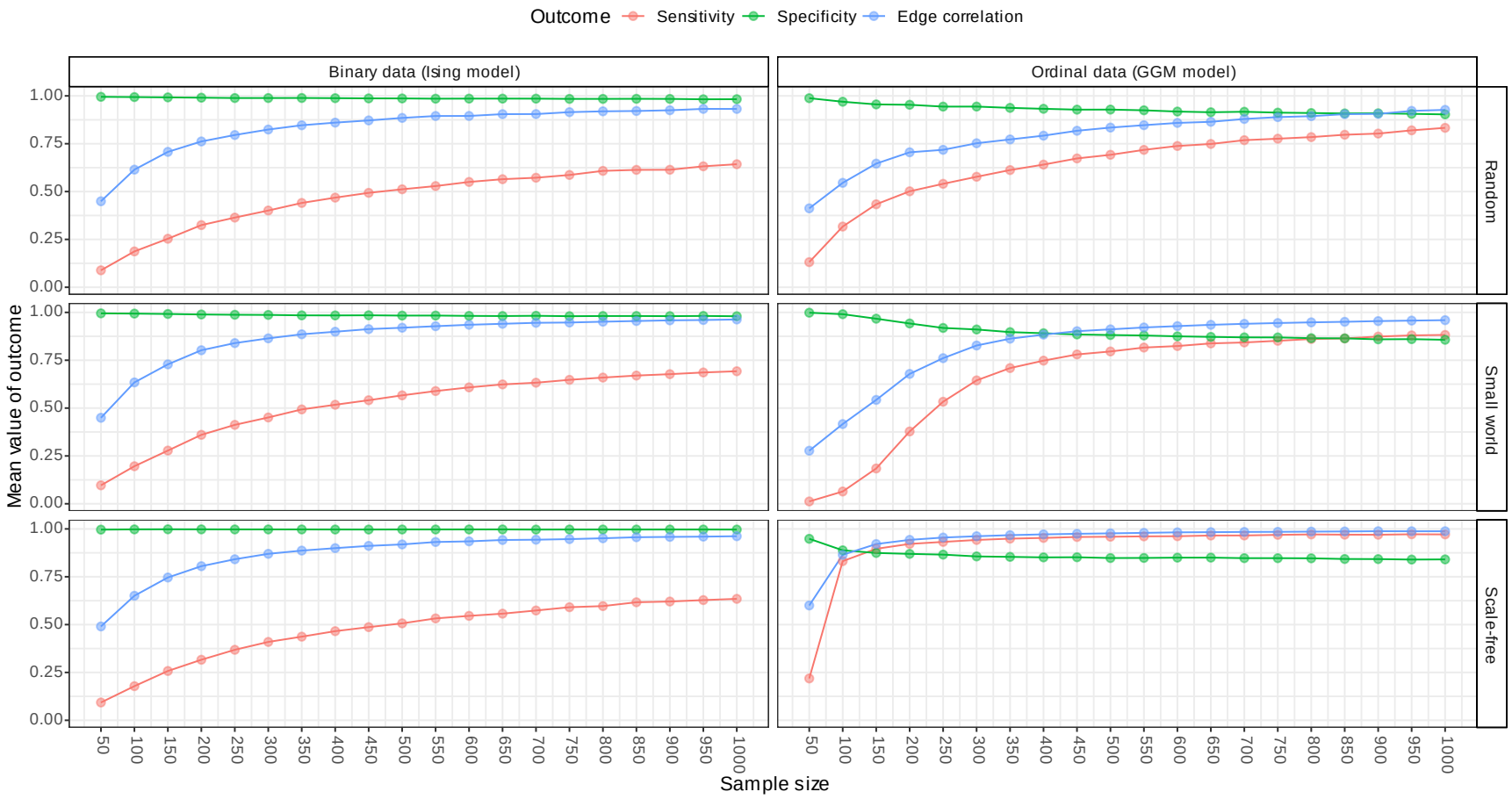

Figure 4. Sensitivity, specificity and edge weights correlation of simulated data. Horizontal panels indicate the type of data used and vertical panels the network architecture. The color of the lines indicates different outcome measures. The edge weights correlation was computed between the upper triangle of the data generating model (i.e., true model) and estimated model.

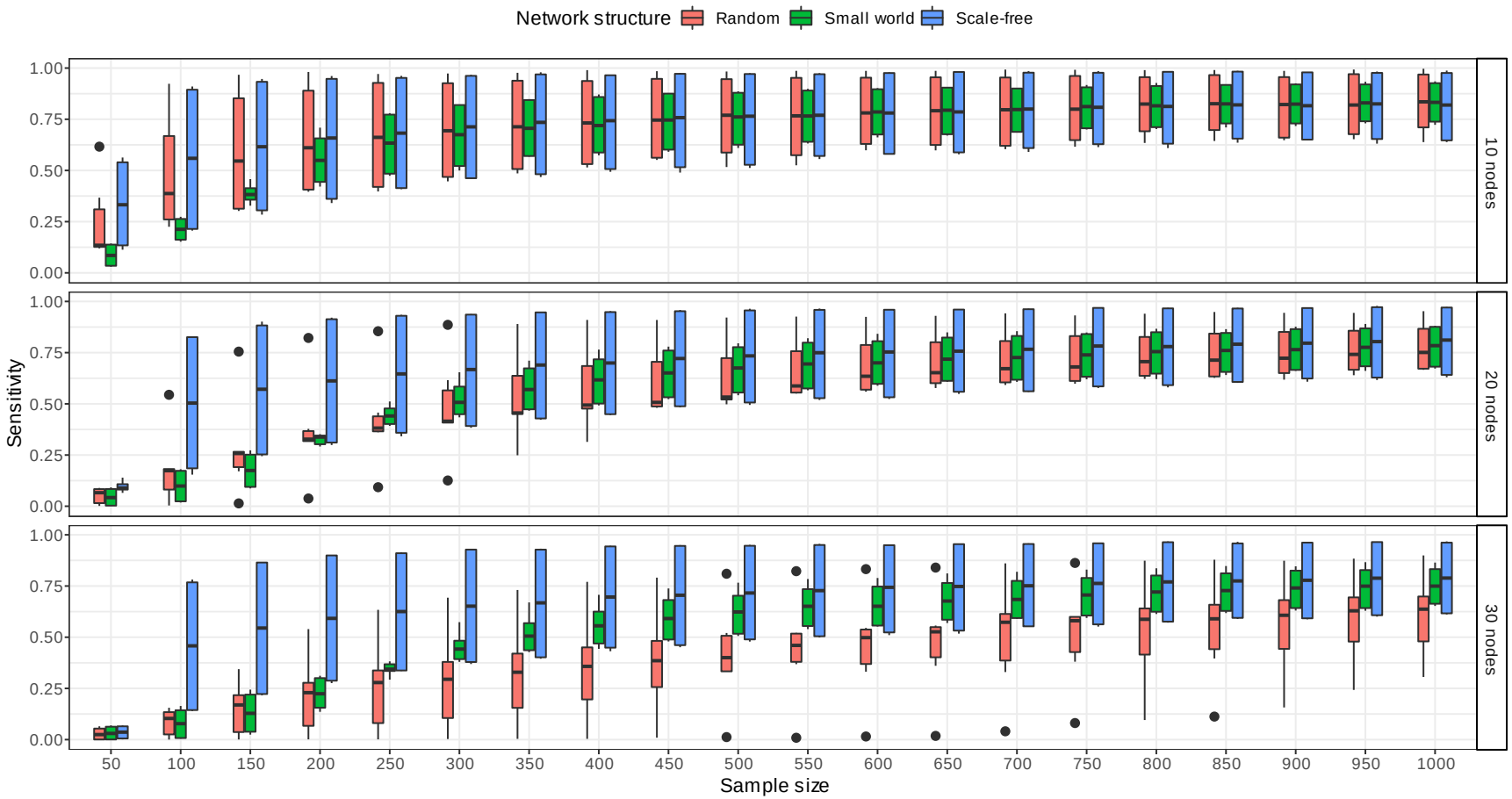

Figure 5. Sensitivity of simulated data presented in Tukey boxplots. Vertical panels indicate the number of nodes in the network structure and the color of the boxplots indicates different network structures.

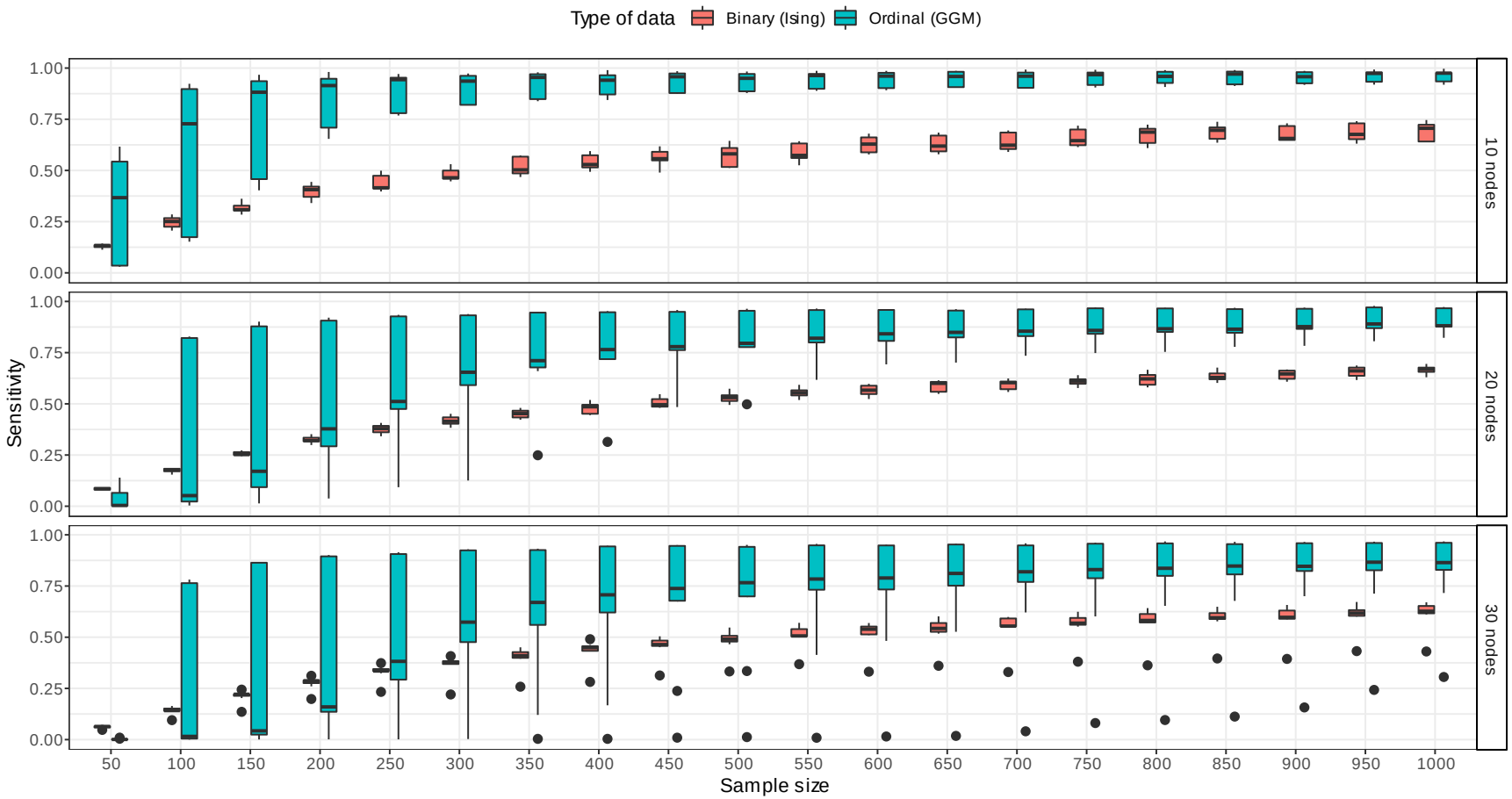

Figure 6. Sensitivity of simulated data presented in Tukey boxplots. Vertical panels indicate the number of nodes in the network structure and the color of the boxplots indicates different types of data and the network models used for estimation.

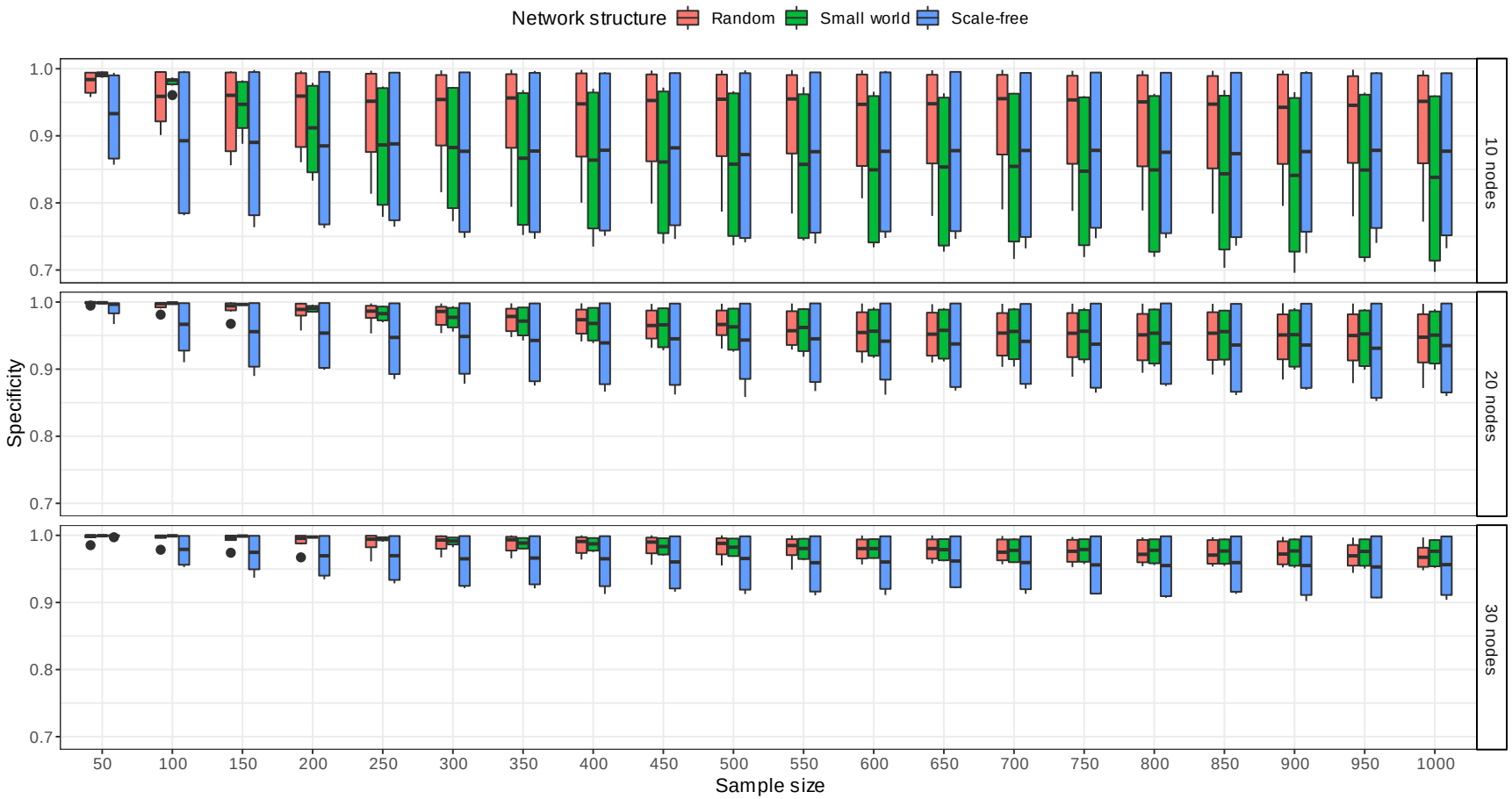

Figure 7. Specificity of simulated data presented in Tukey boxplots. Vertical panels indicate the number of nodes in the network structure and the color of the boxplots indicates different network structures.

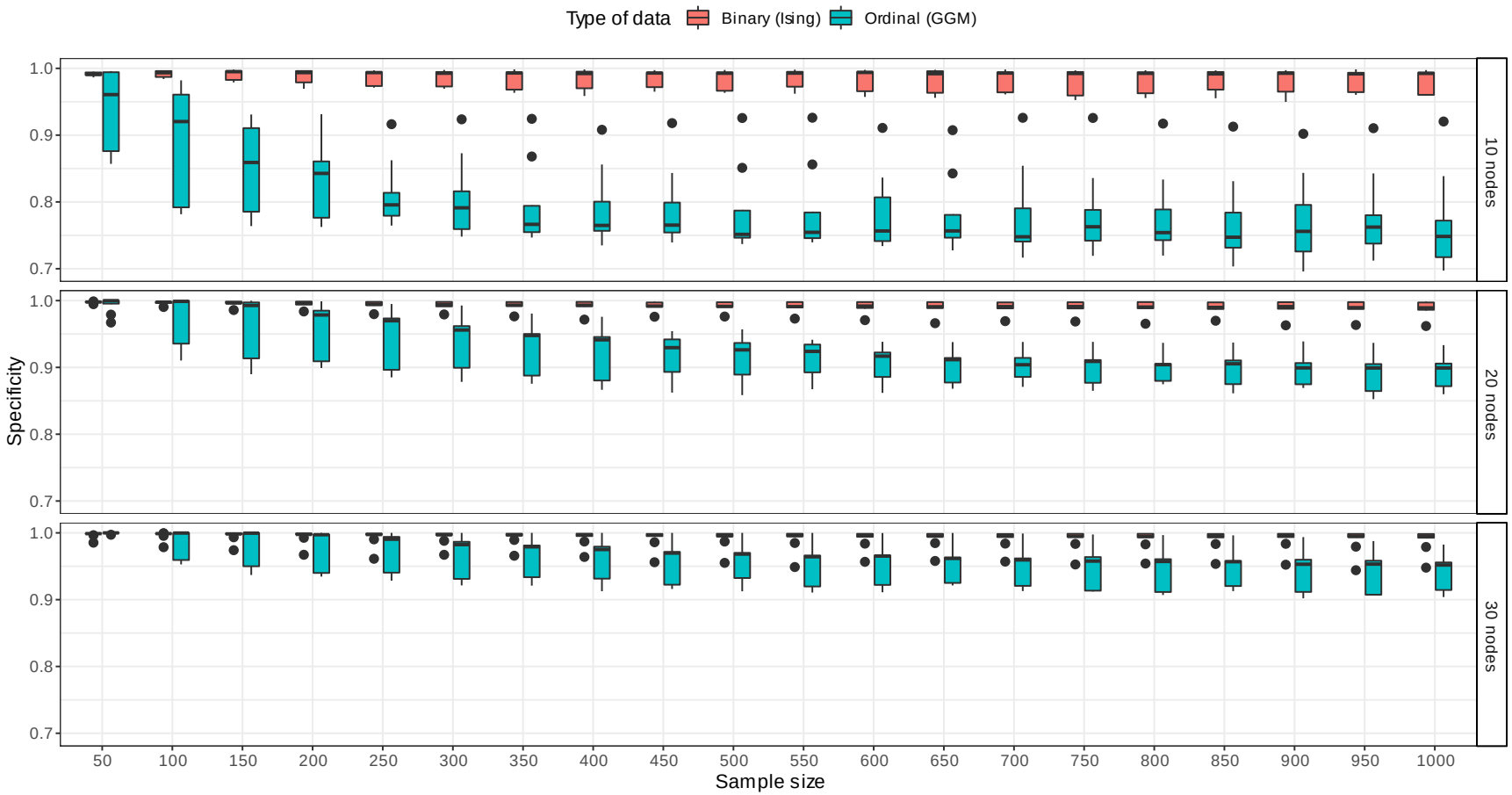

Figure 8. Specificity of simulated data presented in Tukey boxplots. Vertical panels indicate the number of nodes in the network structure and the color of the boxplots indicates different types of data and the network models used for estimation.

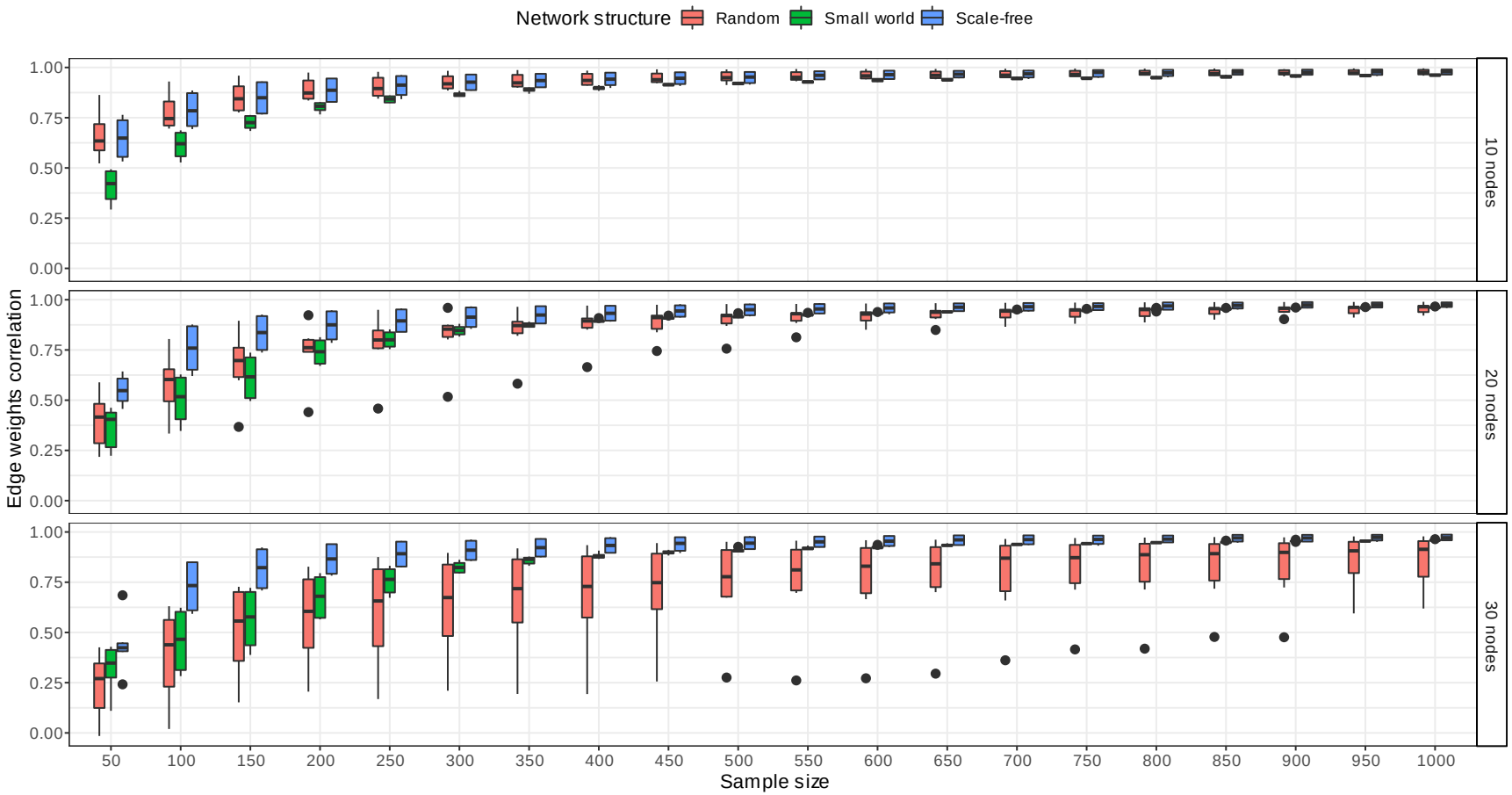

Figure 9. Edge weights correlation between the upper triangle of the data generating model and the estimated model, presented in Tukey boxplots. Vertical panels indicate the number of nodes in the network structure and the color of the boxplots indicates different network structures.