Testing for Stochastic Dominance in Social Networks

Firmin Doko Tchatoka, Robert Garrard and Virginie Masson
School of Economics, University of Adelaide

CRICOS PROVIDER 00123M

adelaide.edu.au

seek LIGHT

Testing for Stochastic Dominance in Social Networks

Firmin Doko Tchatoka*, Robert Garrard, and Virginie Masson

School of Economics, The University of Adelaide

October 8, 2016

Abstract

This paper illustrates how stochastic dominance criteria can be used to rank social networks in terms of efficiency, and develops statistical inference procedures for assessing these criteria. The tests proposed can be viewed as extensions of a Pearson goodness-of-fit test and a studentized maximum modulus test often used to partially rank income distributions and inequality measures. We establish uniform convergence of the empirical size of the tests to the nominal level, and show their consistency under the usual conditions that guarantee the validity of the approximation of a multinomial distribution to a Gaussian distribution. Furthermore, we propose a bootstrap method that enhances the finite-sample properties of the tests. The performance of the tests is illustrated via Monte Carlo experiments and an empirical application to risk sharing networks in rural India.

*Corresponding author. E-mail: firmin.dokotchatoka@adelaide.edu.au. Address: The University of Adelaide, School of Economics, Adelaide, SA 5005, Australia.

1 Introduction

This paper considers the problem of assessing stochastic dominance criteria in network theory. Many economic and social interactions involve network relationships, and the role that networks play in determining economic outcomes– such as trade and exchange of goods in non-centralized markets (e.g., Tesfatsion (1997)), provision of mutual insurance in developing countries (e.g., Fafchamps and Lund (2003)), and job search (e.g., Calvo-Armengol (2004))– is now recognized. Recent statistical and econometric studies in network theory have often focused on the estimation of network relationships,[1] and the identification of peer effects.[2] Statistical methods for understanding how individual incentives to form networks align with social efficiency are yet to be developed.

    This paper illustrates how *stochastic dominance* criteria can be used to rank networks in terms of social efficiency, and proposes a nonparametric procedure for assessing these criteria. Often, standard measures– such as the Gini-coefficient or Lorenz curves– are used to rank income and poverty distributions in terms of social efficiency. However, in addition to being relative measures,[3] two income or poverty distributions such that one second-order statistically dominates the other may result in a same value of these measures. For theses reasons, stochastic dominance criteria are usually preferred to provide a partial ordering of inequality and poverty measures (e.g., Atkinson (1987) and Anderson (1996)), and the concept, as well as its connection to social welfare theory, now extends to network theory (e.g., Goyal (2012) and Jackson et al. (2008)). To illustrate how the stochastic dominance criteria could provide a partial ordering of networks, let $N = \{1, 2, \ldots, n\}$ be a finite set of n agents and $G(N)$ be the set of networks on N. Let $\mathbb{W}(d_g)$ denote the aggregate social welfare function of network $g \in G(N)$, where $d_g = (d_{g \cdot 1}, \ldots, d_{g \cdot n})'$ and $d_{g \cdot i}$ is the degree of agent $i \in N$ in g. Following Goyal (2012, Section 7.4), network $g \in G(N)$ is said to be *socially efficient*

---

[1]See Chandrasekhar (2016), Leung (2015), Banerjee et al. (2013), Liu (2013), Bickel et al. (2011), and Bickel and Chen (2009) among others.

[2]See Hsieh and Lee (2016), Blume et al. (2015), Bursztyn et al. (2014), Goldsmith-Pinkham and Imbens (2013), Jackson (2014), Graham (2014), and Aliprantis and Richter (2013) among others.

[3]For example, changing income inequality, measured by Gini-coefficients, can be due to structural changes in a society such as aging populations, emigration,immigration, etc.

if $\mathbb{W}(d_g) \geqslant \mathbb{W}(d_{g'})$ for all $g' \in G(N)$. Therefore, if $\mathbb{W}(d_g)$ is a nondecreasing and strictly concave function of $d_{g\cdot i}$ for all $i \in N$, then second-order stochastic dominance between the degree distributions of two networks g and g' in $G(N)$ is equivalent to dominance between $\mathbb{W}(d_g)$ and $\mathbb{W}(d_{g'})$ in the same direction (e.g., Rothschild and Stiglitz (1970)). Therefore, the stochastic dominance criteria provide a partial ordering of the elements of $G(N)$ in terms of social efficiency in this setting, and developing statistical methods to establish this ordering from the observed network relationships can be of great interest in social science.

Tests similar to that of Pearson (1900) are often used for assessing stochastic dominance hypotheses in the literature on inequality and poverty measures,[4] but to the best of our knowledge, this study is the first to focus on extending these procedures to network theory. Anderson (1996) suggests a combination of Pearson-type and studentized maximum modulus (SMM) tests[5] in a single decision rule for assessing stochastic dominance of income distributions. His methodology is nonetheless not directly applicable in the context of networks for the following reasons. *First*, both tests are derived in his framework under the assumption that the samples are independent. Although this may be reasonable in the literature on income distributions and poverty measures, it is less likely to be the case in network theory, as it excludes interesting situations where networks' populations overlap. For example, when comparing risk sharing networks formed by men and women within a village (or community), it is reasonable to assume that the two networks are independent across households, while the correlation between the two networks is likely high within households. *Second*, partitioning of samples into classes is usually required to implement a Pearson-type test, and it is well documented that such a partitioning has an influence on the properties (size and power) of the resulting test.[6] In the case where the samples are drawn from a continuous distribution, Mann and Wald (1942) and Williams (1950) propose rules of thumb to select the number of classes and the lengths of subsequent intervals such that the resulting test is unbiased. These optimal rules are usually obtained by equalizing cell probabilities under the

---

[4]For example, see McFadden (1989), Anderson (1996), Davidson and Duclos (2000), Barrett and Donald (2003), Linton et al. (2005), and Barrett et al. (2014).

[5]See Stoline and Ury (1979) for the tabulation of the critical values of the SMM statistics.

[6]See Hotelling (1930), Mann and Wald (1942), Gumbel (1943), Williams (1950), Cochran (1952), and Schorr (1974) among others.

null whilst maintaining an expected cell frequency of at least 5 (e.g., Anderson (1996)). The main difficulty in extending Mann and Wald (1942) and Williams (1950) rules of thumb to the context of networks resides in the finite and discrete nature of the range of a network's degree distribution.

Our contribution in this paper is threefold. *First*, we propose an adjustment to Mann and Wald (1942) and Williams (1950) rules of thumb that applies to the context of networks. We show how the optimal choice of the number of classes can be approximated through a careful analysis of the empirical histogram of the degree distributions of the networks. *Second*, we propose a generalization of the Pearson- and SMM-type statistics in Anderson (1996) that are valid even when the samples are correlated, thus applicable to the context of network theory. Our statistics differ from that of Anderson (1996) and prior literature not only through the correction to account for the correlation between the degree distributions of the networks, but also their direct dependence on partitioning into classes. We show that a combination of the two modified statistics into a single decision rule is necessary to inform us on whether *stochastic dominance* holds or not, once equality between the degree distributions of the networks is rejected. As the modified statistics depend on partitioning into classes, controlling the size of the resulting tests uniformly over the set of all *admissible partitions*[7] is important for the asymptotic results to give a good approximation of the empirical size to the nominal level. *Finally*, we provide a bootstrap procedure that improves the finite-sample performance of both the modified Pearson- and SMM-statistics.

We provide an analysis of both the size and power properties of the tests under weaker assumptions than is usually the case in most applications of Pearson's (1900) goodness-of-fit test. On level control, we establish uniform convergence of their empirical size to the nominal level over the set of all *admissible partitions* when the usual asymptotic chi-square and SMM critical values are applied. On power, we show that test consistency holds no matter which *admissible* partition is used. Moreover, we establish uniform consistency of the bootstrap for the two modified Pearson- and SMM-tests irrespective of whether the null hypothesis holds or not. We present a Monte Carlo experiment that confirms our theoretical findings. In particular, while the standard tests sometimes tend to over-reject the null hypothesis

---

[7]An *admissible partition* is a partition in which the minimum expected number in each cell is at least 5.

if the sample size is small, the bootstrap tests have an overall good performance in such contexts. Finally, using the data set of Jackson et al. (2012) and Banerjee et al. (2012, 2013), we illustrate our theory through an investigation of the households' *risk sharing networks* across 75 villages in rural India. In particular, we focus on both the goods lending and money lending networks, and test gender differences within these networks by applying the tests of stochastic dominance developed. For goods lending, both the standard and bootstrap tests show that the *female network* first- and second-order stochastically dominates the *male network* at the 1% and 5% nominal levels. However, for money lending, we could only find evidence of the first- and second-order dominance of the *female network* at the 5% nominal level. At the 1% nominal level, neither network dominates the other with both the standard and bootstrap tests. These results suggest that women within these villages overall tend to form denser *risk sharing networks* than do men, especially for goods lending.

Throughout this paper, for any vector $x = (x_1, \ldots, x_k)' \in \mathbb{R}^k$, the notation "$x \leqslant 0$" means $x_l \leqslant 0$ for all $l = 1, \ldots, k$, while "$x \nleqslant 0$"(or "$x \ngeqslant 0$") means that there exists l and l' in $\{1, \ldots, k\}$ such that $x_l \geqslant 0$ and $x_{l'} < 0$ or $x_l > 0$ and $x_{l'} \leqslant 0$. Convergence almost surely is symbolized by "*a.s.*", "$\overset{p}{\to}$" stands for convergence in probability, while "$\overset{d}{\to}$" means convergence in distribution. The usual stochastic orders of magnitude are denoted by $O_p(.)$, $o_p(.)$. $\mathbb{P}[\cdot]$ denotes the relevant probability measure and $\mathbb{E}[\cdot]$ is the expectation operator under $\mathbb{P}[\cdot]$. $\mathbb{P}^*[\cdot]$ is the bootstrap analogue of $\mathbb{P}[\cdot]$, and similarly for $\mathbb{E}^*[\cdot]$. I_q stands for the identity matrix of order q, and for any $q \times q$ matrix A, A^- is the generalized inverse of A. The notation $\mathbf{diag}(A)$ is a $q \times q$ diagonal matrix with diagonal elements the $(l, l)^{th}$ elements of A. $\|U\|$ denotes the usual Euclidian or Frobenius norm for a matrix U. For any set $\mathscr{C}$, $\partial \mathscr{C}$ is the boundary of $\mathscr{C}$ and $(\partial \mathscr{C})^\epsilon$ its ϵ-neighborhood. Finally, $\sup_{\omega \in \Omega} |f(\omega)|$ is the supremum norm on the space of bounded continuous real functions, with topological space Ω.

The remainder of the paper is organised as follows. Section 2 defines the relevant concepts and introduces the dominance criterion. Section 3 formulates the hypotheses tested and presents the basic notations and assumptions used. Section 4 presents the derivation of the statistics and the asymptotic theory developed. Section 5 illustrates the performance of the tests via Monte Carlo experiments. Section 6 provides an empirical illustration of our theoretical results, and Section 7 concludes. Proofs are presented in the appendix.

2 Preliminaries

Before introducing the concept of stochastic dominance in networks (Section 2.2) and formalizing the testing problem of interest (Section 3), we define the basic terminologies and notations used throughout the study.

2.1 Networks

Let $N = \{1, 2, \ldots, n\}$ denote a finite set of agents, and $G(N)$ be the set of networks on N. We define a *network* g over N as a pair of nodes and edges describing relationships (or links) between agents $1, 2, \ldots, n$. A network can be represented by a graph whose $n \times n$ adjacency matrix has generic element $g_{ii'}$ satisfying $g_{ii'} = 1$ if there is a directed link from agent i to i', and $g_{ii'} = 0$ otherwise. By convention, we set $g_{ii} = 0$ for all i. The *neighborhood* of agent i is the set of agents with whom i has a directed link in network g, i.e., the set $\mathcal{N}_i(g) = \{i' \in N | g_{ii'} = 1\}$. We refer to the number of agent i's neighbors, $d_{g \cdot i} = card[\mathcal{N}_i(g)]$, as the *degree* of agent i.[8]

The *degree distribution* of network g is a vector $P_g = [\hat{p}_{g \cdot 0}, \ldots, \hat{p}_{g \cdot k}, \ldots, \hat{p}_{g \cdot (n-1)}]'$, where $\hat{p}_{g \cdot k} = card[\{i : d_{g \cdot i} = k\}]/n$ is the proportion of nodes with degree k; thus $\hat{p}_{g \cdot k} \geqslant 0$ for each $k \in \mathcal{R}_n$, $\sum_{k \in \mathcal{R}_n} \hat{p}_{g \cdot k} = 1$, and $\mathcal{R}_n = \{0, 1, \ldots, n-1\}$ is the range of P_g. The empirical *cumulative distribution function* (cdf) of network g is the function $F_g : \mathcal{R}_n \to [0, 1]$ such that $F_g(k) = \sum_0^k \hat{p}_{g \cdot l}$ for all $k \in \mathcal{R}_n$.

Example 1. Figure 1 illustrates three networks with $n = 5$ agents: a "circle" network (Network g), a "directed star" network (Network g'), and a "complete" network (Network g'').

---

[8]Our definition of a neighborhood considers the *out-degree* of agent i, i.e. the number of links which originate from agent i. However, it can also be defined using the *in-degree* of agent i, in which case, $\mathcal{N}_i(g) = \{i' \in N | g_{i'i} = 1\}$. The choice of the definition depends mainly upon the application considered. For undirected networks, $g_{ii'} = g_{i'i}$ and both definitions coincide.

Figure 1: Example of networks

(a) Network g (b) Network g' (c) Network g''
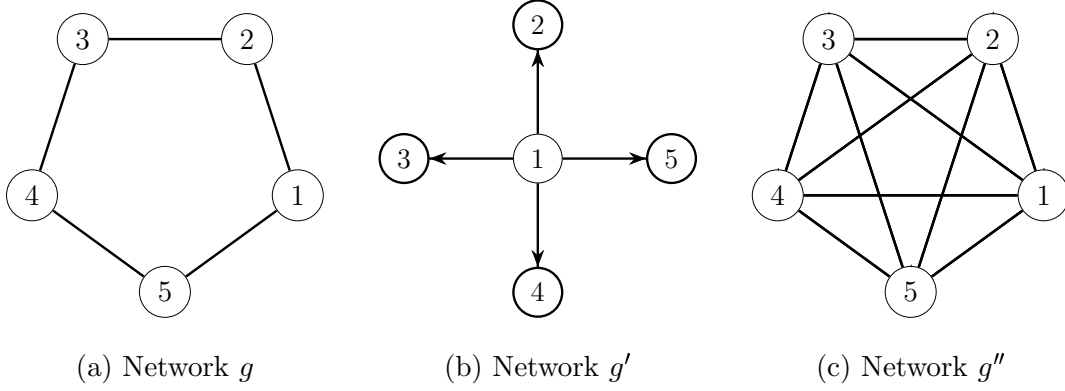
The characteristics of each network $j \in \{g, g', g''\}$, as per the above terminologies and definitions– *neighborhood*: $\mathscr{N}(j)$, *degree* of agent: $d_{j \cdot i}$, *degree distribution*: P_j, and *empirical cdf*: F_j –are summarized in Table 1.

Table 1: Characteristics of network $j \in \{g, g', g''\}$

characteristics $\downarrow$ Network j $\rightarrow$	g	g'	g''
$\mathscr{N}_1(j)$	$\{2,5\}$	$\{2,3,4,5\}$	$\{2,3,4,5\}$
$\mathscr{N}_2(j)$	$\{1,3\}$	$\emptyset$	$\{1,3,4,5\}$
$\mathscr{N}_3(j)$	$\{2,4\}$	$\emptyset$	$\{1,2,4,5\}$
$\mathscr{N}_4(j)$	$\{3,5\}$	$\emptyset$	$\{1,2,3,5\}$
$\mathscr{N}_5(j)$	$\{1,4\}$	$\emptyset$	$\{1,2,3,4\}$
$d_{j \cdot 1}$	2	4	4
$d_{j \cdot 2}$	2	0	4
$d_{j \cdot 3}$	2	0	4
$d_{j \cdot 4}$	2	0	4
$d_{j \cdot 5}$	2	0	4
P_j	$(0,0,1,0,0)'$	$(4/5,0,0,0,1/5)'$	$(0,0,0,0,1)'$
F_j	$(0,0,1,1,1)'$	$(4/5,4/5,4/5,4/5,1)'$	$(0,0,0,0,1)'$

2.2 Stochastic Dominance in Networks

Consider the setup described in Section 2.1, and let g and g' denote two networks in $G(N)$ with empirical cdfs F_g and $F_{g'}$, respectively. The first- and second-order[9] stochastic dominance between g and g' are characterized as follows.

Definition 1. (**i**) Network g **first-order stochastically dominates** network g', which we write $g \succ_1 g'$, if $F_g(k) \leqslant F_{g'}(k) \; \forall \; k \in \mathscr{R}_n$, with strict inequality for some k.
(**ii**) Network g **second-order stochastically dominates** g', which we write $g \succ_2 g'$, if $\sum_{i=0}^{k} F_g(i) \leqslant \sum_{i=0}^{k} F_{g'}(i) \; \forall \; k \in \mathscr{R}_n$, with strict inequality for some k.

It is straightforward to see from the above characterizations that first-order stochastic dominance implies second-order stochastic dominance, but not the other way around. We now illustrate the two concepts from the example of Section 2.1.

Example 1 (continued)**.** Again, consider the three networks g, g', and g'' of Example 1. From Table 2 below, the pairwise comparisons between the cumulative distributions of these networks show that g'' first-order stochastically dominates both g and g'. Therefore, g'' also second-order stochastically dominates both g and g'. However, as $F_g(1) < F_{g'}(1)$ and $F_g(2) > F_{g'}(2)$, there exists no first-order stochastic dominance between g and g'. Nevertheless, g second-order stochastically dominates g'. This reflects the fact that network g has an average degree at least as high as network g' but a lower dispersion in agents' degrees.

---

[9]The characterization of stochastic dominance can easily be extended to higher-order, but for simplicity we mainly focus on the first- and second-order dominance for the remainder of the paper.

Table 2: Stochastic dominance between networks g, g' and g'' of Example 1

k	0	1	2	3	4
$\hat{p}_{g \cdot k}$	0	0	1	0	0
$F_g(k)$	0	0	1	1	1
$\sum_{i=0}^{k} F_g(i)$	0	0	1	2	3
$\hat{p}_{g' \cdot k}$	0.8	0	0	0	0.2
$F_{g'}(k)$	0.8	0.8	0.8	0.8	1
$\sum_{i=0}^{k} F_{g'}(i)$	0.8	1.6	2.4	3.2	4.2
$\hat{p}_{g'' \cdot k}$	0	0	0	0	1
$F_{g''}(k)$	0	0	0	0	1
$\sum_{i=0}^{k} F_{g''}(i)$	0	0	0	0	1

We now wish to formulate hypotheses for assessing stochastic dominance in social networks from observed real world data.

3 Stochastic Dominance Hypothesis and Assumptions

We first formulate the problem of testing stochastic dominance hypotheses in Section 3.1. Section 3.2 presents the basic notations and assumptions that are used in the paper.

3.1 Hypothesis Formulation

Let g and g' be two networks observed on the same population of n agents, and let F_j denote the empirical cdf associated with the degree distribution P_j of network $j \in \{g,\ g'\}$. Finally, let $\mathbb{N} = \{1, 2, \ldots\}$ be the set of natural integers. Given $m \in \mathbb{N}$, we are interested in assessing which network mth-order *stochastically dominates* the other. From Definition 1, this problem can be formulated as a problem of testing the mth-order stochastic dominance between the cdfs F_g and $F_{g'}$, i.e.,

$$H_{0m} : F_g \stackrel{d}{=} F_{g'} \text{ versus } H_{1m} : F_g \succ_m F_{g'} \ \wedge \ H_{2m} : F_g \stackrel{d}{\neq} F_{g'}, \tag{1}$$

where "$\succ_m$" denotes the mth-order stochastic dominance operator, "$\overset{d}{=}$" and "$\overset{d}{\neq}$" symbolize equality and difference in distribution respectively. As can be seen clearly from (1), H_{0m} tests equality between F_g and $F_{g'}$ against: (i) mth-order stochastic dominance (H_{1m}), and (ii) no mth-order dominance (H_{2m}). For example when $m = 2$, H_{02} tests the equality between F_g and $F_{g'}$ against both second-order stochastic dominance (H_{12}) and no second-order dominance (H_{22}). Several statistical procedures exist to assess stochastic dominance hypotheses between two distributions, but to the best of our knowledge, this study is the first to focus on extending these procedures to network theory.

In order to derive a testable formulation of problem (1) from the observed data, as well as test statistics for assessing it, it is useful to first introduce the following notations and assumptions.

3.2 Basic Notations and Assumptions

Let $\{(d_{g \cdot i},\, d_{g' \cdot i})\}_{i=1}^{n}$ be a sample of n observations drawn from the joint distribution of the *degree of agents* in networks g and g'. Let F_g and $F_{g'}$ denote the empirical cdfs of networks g and g' respectively, constructed as in Section 2.1. To build Pearson-type statistics for assessing H_{0m} in (1), we must first partition the range (support) of the *degree distributions* of networks g and g' into classes (or class intervals). To do this, we adapt the methodology in Anderson (1996) to the context of social networks.

Let $(d_i)_{i=1}^{2n}$ be the pooled sample of $2n$ observations obtained by stacking the two sub-samples $(d_{g \cdot i})_{i=1}^{n}$ and $(d_{g' \cdot i})_{i=1}^{n}$, and let $Supp(d) \subseteq \mathscr{R}_n$ denote the support of the distribution of $(d_i)_{i=1}^{2n}$, where $\mathscr{R}_n = \{0, 1, 2, \ldots, n-1\}$ is the common range of the *degree distributions* of networks g and g'. Note that $Supp(d)$ need not be strictly equal to $\mathscr{R}_n$. This is the case for example if $\max_{i,j \in \{g,g'\}} \{d_{j \cdot i}\}_{i=1}^{n} < n - 1$. For some fixed $k \in \mathbb{N}$, let $\mathbf{P}_n^{(k)}(\mathbf{I}_1, \ldots, \mathbf{I}_k) \equiv \mathbf{P}_n^{(k)}(\mathbf{I}) :=$ $\{\mathbf{I}_l\}_{l=1}^{k}$ denote a finite partition of $Supp(d)$ into k disjoint sets, i.e.

$$Supp(d) = \bigcup_{1 \leqslant l \leqslant k} \mathbf{I}_l : \ \mathbf{I}_l \neq \emptyset, \ \mathbf{I}_l \cap \mathbf{I}_{\tilde{l}} = \emptyset \ \forall \, l \neq \tilde{l}, \tag{2}$$

and define a collection of such partitions by

$$\mathscr{P} = \left\{ \mathbf{P}_n^{(k)}(\mathbf{I}) : \mathbf{I} = \{\mathbf{I}_l\}_{l=1}^{k} \text{ satisfies (2)} \right\}. \tag{3}$$

10

As *Supp(d)* is a discrete finite set, the collection $\mathscr{P}$ contains a finite number of elements (or partitions) for a given k. Until now, we have implicitly assumed that the number k of subsets and the division points between subsets (subsets' cardinality) in (2) are available to the investigator. In practice, one has to choose k as well as the division points between the k resulting subsets, and it is well documented that these choices have an influence on the properties (size and power) of Pearson-type tests. For samples generated from continuous distributions, we have $Supp(d) \subseteq \mathbb{R}$ and $\mathbf{I}_l$, $l = 1, 2, \ldots, k$ are compact intervals in (2). In this case, there is a number of seminal papers which provide rules to select k and the lengths of subsequent intervals such that the resulting Pearson-type test is unbiased. For example, Anderson (1996) suggests that power can be gained by locating partition points at fractiles where it is thought that the two distributions may intersect. Since this information is unknown, the standard advice by Mann and Wald (1942),[10] Gumbel (1943), and Williams (1950), that power is gained by equalizing cell probabilities under the null whilst maintaining an expected cell frequency of at least 5 is usually used in applied work.

The main difficulty in extending Mann and Wald (1942) and Williams (1950) rules of thumb to the context of networks resides in the finite and discrete nature of the range of a network's degree distribution. For example, Figure 2 shows the degree distributions of two commonly used networks: the *Poisson random graph* and the *Scale-free network*. While in theory the range of both distributions is the entire positive integer set $\mathbb{N}$, we see that both distributions are concentrated between: 1–20 (for the *Poisson random graph*), and 1–9 (the *Scale-free network*). Suppose we have a joint sample of $n = 500$ realizations of networks g and g' drawn from a population that follows one of these distributions. For a test at the $\alpha = 5\%$ nominal level ($c = 1.64$), Mann and Wald's (1942) and Williams's (1950) optimal

---

[10]Mann and Wald (1942) show that the optimal choice of the number of classes is $k := \mathscr{I}nt\left[4\sqrt[5]{\frac{2(n-1)^2}{c^2}}\right]$, where n is the sample size, $\mathscr{I}nt[x]$ is the integer part of any real x, and c is determined so that $\frac{2}{\sqrt{2\pi}}\int_c^\infty e^{-x^2/2}dx$ is equal to the size of the critical region under H_{02}. One criticism of Mann and Wald's (1942) method is that it generates an unnecessarily large number of classes; see Schorr (1974). Williams (1950) shows that halving this number does not substantially decrease the power of Pearson-type tests. Although these rules of thumb are reasonable to follow, it is worth noting that they do not imply that the resulting Pearson-type test is necessarily uniformly powerful against all alternatives; for example, see Cochran (1952).

rules of thumb give $k_{MW} = 45$ and $k_W = 23$ respectively. These choices increase to $k_{MW} = 59$ and $k_W = 30$ for a population of $n = 1,000$ agents. However, Figure 2 shows clearly that even a choice of $k \equiv k_W = 23$ in (2) does not make it possible to equalize cell probabilities under the null whilst maintaining an expected cell frequency of at least 5. Even though this criterion may give a good approximation for *Poisson random graphs* in some instances (for example when λ is large enough), this is likely not the case for *Scale-free networks*. Therefore, adjustments are needed to adapt Mann and Wald's (1942) and Williams's (1950) rules of thumb to the network context. For this purpose, define $k_{max} = \max Supp(d)$. Then, a practical and simple rule of thumb could be to choose $k \leqslant \min[k_W, \max supp(d)]$ such that Williams (1950) rule of thumb is close to being fulfilled. This can be achieved through a careful analysis of the empirical histogram of the degree distributions such as in Figure 2. For example, if the realizations of networks g and g' are drawn from a *Poisson* population (Figure 2-(a)), both choices: (i) $k = 4$ and $\mathbf{I}_1 = \{1, \ldots, 7\}$, $\mathbf{I}_2 = \{8, 9\}$, $\mathbf{I}_3 = \{10, 11\}$, $\mathbf{I}_4 = \{12, \ldots, 20\}$, and (ii) $k = 4$ and $\mathbf{I}_1 = \{1, \ldots, 9\}$, $\mathbf{I}_2 = \{10\}$, $\mathbf{I}_3 = \{11\}$, $\mathbf{I}_4 = \{12, \ldots, 20\}$, are acceptable. However, the former is closer to the recommendation to equalize cell probabilities than the latter.

Figure 2: The distribution of degrees for Poisson and Scale-free networks

(a) Poisson with parameter $\lambda = 10$ (b) Scale-free with parameter $\gamma = 2.5$
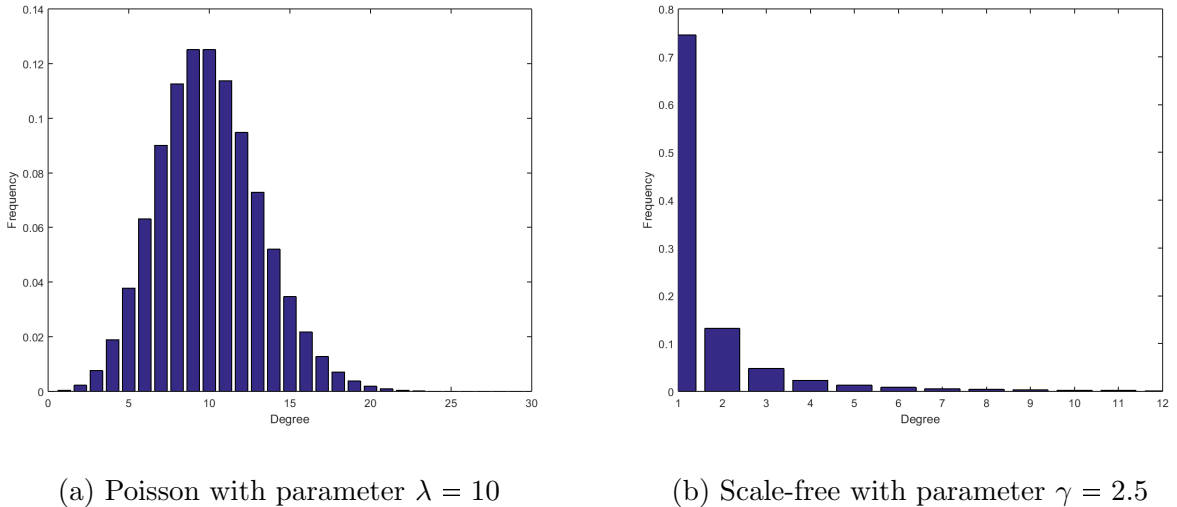
To formally address the threshold of an expected cell frequency of at least 5, we first

introduce the following notations and definitions. Let $p_{j \cdot il}, j \in \{g, g'\}$ be the probability that $d_{j \cdot i}$ falls in $\mathbf{I}_l$, and $\hat{p}_{j \cdot l}$ denote the proportion of observations in $(d_{j \cdot i})_{i=1}^n$ which fall in $\mathbf{I}_l$, i.e.

$$p_{j \cdot il} = \mathbb{P}(d_{j \cdot i} \in \mathbf{I}_l), \quad \hat{p}_{j \cdot l} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(d_{j \cdot i} \in \mathbf{I}_l). \tag{4}$$

If $\{(d_{g \cdot i}, d_{g' \cdot i})\}_{i=1}^n$ is i.i.d. across i, for given $j \in \{g, g'\}$ and $l \in \{1, \ldots, k\}$, the probabilities $p_{j \cdot il}$ are the same for all i, i.e., $p_{j \cdot il} \equiv p_{j \cdot l}$ for all i and $\hat{p}_{j \cdot l}$ is a consistent estimator of $p_{j \cdot l}$. Then, the *expected numbers* in cell l for network j is given by

$$n_{j \cdot l} := n\hat{p}_{j \cdot l} = \sum_{i=1}^n \mathbb{1}(d_{j \cdot i} \in \mathbf{I}_l). \tag{5}$$

To insure a valid approximation of the *multinomial distribution* to a *multivariate normal distribution*, (2) must also guarantee that the minimum of the $n_{j \cdot l}$'s for all $j \in \{g, g'\}$ and $l = 1, \ldots, k$ is at least 5. This threshold is usually imposed and the absence of a theory to justify its validity has raised some concerns in several seminal papers; e.g., Cochran (1952), Lewis and Burke (1949), and Edwards (1950). Yates (1934) provides a correction for continuity that adjusts the formula for a Pearson-type statistic when this threshold is violated. In this paper, we do not address the issues related to the choice of the minimum expected number in cells. Rather, we consider the collection of all partitions $\mathbf{P}_n^{(k)}(\mathbf{I})$ for which this requirement is satisfied, and we wish to provide tests of stochastic dominance that control the size uniformly over this collection of partitions.

To be more specific, consider the partitions $\mathbf{P}_n^{(k)}(\mathbf{I})$ in (2) such that $n\hat{p}_{j \cdot l} > 5$ for all $j \in \{g, g'\}$ and $l \in \{1, \ldots, k\}$. Let $\mathscr{P}_A$ be a collection of such partitions, i.e.

$$\mathscr{P}_A = \left\{ \mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P} : \mathbf{I} = \{\mathbf{I}_l\}_{l=1}^k \text{ satisfies } n\hat{p}_{j \cdot l} > 5; \text{ for all } j \in \{g, g'\} \text{ and } l = 1, \ldots, k \right\}. \tag{6}$$

For the remainder of the paper, we shall refer to $\mathscr{P}_A$ as a collection of *admissible partitions*. Note that $n\hat{p}_{j \cdot l} > 5$ is the only restriction on the structure of $\mathbf{P}_n^{(k)}(\mathbf{I})$ in (2), therefore there are many admissible partitions $\mathbf{P}_n^{(k)}(\mathbf{I})$ that can be formed from the observed joint data $\{(d_{g \cdot i}, d_{g' \cdot i})\}_{i=1}^n$. As $\mathscr{P}$ is finite, $\mathscr{P}_A$ is also a finite set of partitions. In such a context, proving the uniform control of *type-I error* over $\mathscr{P}_A$ of the statistics considered for assessing H_{0m} in (1) is important.

13

Now, let

$$u_{j \cdot i} \;=\; [\mathbb{1}(d_{j \cdot i} \in \mathbf{I}_1), \;\ldots, \;\mathbb{1}(d_{j \cdot i} \in \mathbf{I}_k)]', \quad p_{j \cdot i} = \mathbb{E}(u_{j \cdot i}) \equiv [p_{j \cdot i1}, \ldots, p_{j \cdot ik}]', \quad (7)$$

$$\text{and } \hat{p}_j \;:=\; [\hat{p}_{j \cdot 1}, \ldots, \hat{p}_{j \cdot k}]' = \frac{1}{n} \sum_{i=1}^{n} u_{j \cdot i}, \; j \in \{g, g'\}, \quad (8)$$

where $p_{j \cdot il}$ and $\hat{p}_{j \cdot l}$ are given in (4). Each estimated vector of probabilities $\hat{p}_j$ in (8) is a *sample average* of the realizations $u_{j \cdot i}$ from a k-dimensional multinomial random variable with vector of parameters $p_{j \cdot i} = [p_{j \cdot i1}, \ldots, p_{j \cdot ik}]'$. Let $\widehat{\Sigma}_j$ be an estimator of the covariance matrix of $u_{j \cdot i}$ given by

$$\widehat{\Sigma}_j \;=\; \begin{pmatrix} \hat{p}_{j \cdot 1}(1 - \hat{p}_{j \cdot 1}) & -\hat{p}_{j \cdot 1}\hat{p}_{j \cdot 2} & \cdots & -\hat{p}_{j \cdot 1}\hat{p}_{j \cdot k} \\ -\hat{p}_{j \cdot 2}\hat{p}_{j \cdot 1} & \hat{p}_{j \cdot 2}(1 - \hat{p}_{j \cdot 2}) & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ -\hat{p}_{j \cdot k}\hat{p}_{j \cdot 1} & -\hat{p}_{j \cdot k}\hat{p}_{j \cdot 2} & \cdots & \hat{p}_{j \cdot k}(1 - \hat{p}_{j \cdot k}) \end{pmatrix} : \; j \in \{g, g'\}, \quad (9)$$

and similarly, define

$$p_{gg' \cdot i, l\tilde{l}} = \mathbb{P}(d_{g \cdot i} \in \mathbf{I}_l, d_{g' \cdot i} \in \mathbf{I}_{\tilde{l}}), \;\; \hat{p}_{gg' \cdot l\tilde{l}} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(d_{g \cdot i} \in \mathbf{I}_l)\mathbb{1}(d_{g' \cdot i} \in \mathbf{I}_{\tilde{l}}), \quad (10)$$

and let $\widehat{\Sigma}_{gg'}$ be an estimator of the covariance matrix of the $(2k)$-dimensional vector of joint variables $(u'_{g \cdot i} : u'_{g' \cdot i})'$ given by

$$\widehat{\Sigma}_{gg'} \;=\; \begin{pmatrix} \hat{p}_{gg' \cdot 11} - \hat{p}_{g \cdot 1}\hat{p}_{g' \cdot 1} & \hat{p}_{gg' \cdot 12} - \hat{p}_{g \cdot 1}\hat{p}_{g' \cdot 2} & \cdots & \hat{p}_{gg' \cdot 1k} - \hat{p}_{g \cdot 1}\hat{p}_{g' \cdot k} \\ \hat{p}_{gg' \cdot 21} - \hat{p}_{g \cdot 2}\hat{p}_{g' \cdot 1} & \hat{p}_{gg' \cdot 22} - \hat{p}_{g \cdot 2}\hat{p}_{g' \cdot 2} & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \hat{p}_{gg' \cdot k1} - \hat{p}_{g \cdot k}\hat{p}_{g' \cdot 1} & \hat{p}_{gg' \cdot k2} - \hat{p}_{g \cdot k}\hat{p}_{g' \cdot 2} & \cdots & \hat{p}_{gg' \cdot kk} - \hat{p}_{g \cdot k}\hat{p}_{g' \cdot k} \end{pmatrix}. \quad (11)$$

Also, let $\hat{v}_m = \mathbf{T}^m(\hat{p}_g - \hat{p}_{g'})$ be the scaled vector of contrasts, where $\mathbf{T}$ is a $k \times k$ lower triangular matrix of ones, and define

$$\widehat{\Omega}_m = \mathbf{T}^m[\widehat{\Sigma}_g + \widehat{\Sigma}_{g'} - (\widehat{\Sigma}_{gg'} + \widehat{\Sigma}'_{gg'})]\mathbf{T}'^m := [\hat{\omega}_{m \cdot l\tilde{l}}]_{1 \leqslant l, \tilde{l} \leqslant k}. \quad (12)$$

Note that by construction, each of the $k \times k$ matrices $\widehat{\Sigma}_j, j \in \{g, g'\}$ in (9), $\widehat{\Sigma}_{gg'}$ in (11), and $\widehat{\Sigma}_g + \widehat{\Sigma}_{g'} - \widehat{\Sigma}_{gg'} - \widehat{\Sigma}'_{gg'}$ in (12) have rank $k - 1$. Therefore, $\widehat{\Omega}_m$ in (12) also has rank $k - 1$.

14

The notation $\widehat{\Omega}_m^-$ thus refers to the *generalized inverse* of $\widehat{\Omega}_m$ hereinafter. From Dhrymes (1978, Proposition 3.5), there exists a diagonal matrix $\widehat{D}_{k-1}$ whose diagonal elements are the nonzero eigenvalues of $\widehat{\Omega}_m$ (in decreasing order of magnitude), and a $k \times (k-1)$ matrix $\widehat{P}_{k-1}$ whose columns are the (orthogonal) eigenvectors corresponding to the nonzero roots of $\widehat{\Omega}_m$, such that

$$\widehat{\Omega}_m^- = \widehat{P}_{k-1} \widehat{D}_{k-1}^{-1} \widehat{P}_{k-1}'. \tag{13}$$

We now make the following assumption on the joint sample $\{(d_{g \cdot i}, d_{g' \cdot i})\}_{i=1}^n$.

Assumption 1. $\mathscr{D}_n := \{(d_{g \cdot i}, d_{g' \cdot i})\}_{i=1}^n$ *is a i.i.d. random sample across i drawn from the joint distribution of the degrees of networks g and g'.*

In the above assumption, possible dependence between the distribution of the degrees of the two networks is allowed. The i.i.d. sampling across the rows of the joint sample $\mathscr{D}_n$ preserves this dependence. In the case where g and g' are independent, one can draw two independent i.i.d. samples with different sizes: one from the population of network g, say $(d_{g \cdot i})_{i=1}^{n_g}$, and the second from the population of network g', say $(d_{g' \cdot i})_{i=1}^{n_{g'}}$. However, this case excludes interesting situations where the populations of the two networks overlap, as is usually the case in most empirical applications of social networks. In such contexts, while it is reasonable to assume that $(d_{g \cdot i}, d_{g' \cdot i})$ is independent of $(d_{g \cdot i'}, d_{g' \cdot i'})$ for $i \neq i'$, it is likely that $d_{g \cdot i}$ and $d_{g' \cdot i}$ will be correlated.

4   Test Statistics and asymptotic theory

We wish to first discuss how problem (1) can be recast in the more familiar language of hypotheses specified on vectors of contrast. Under the i.i.d. sampling across observations in Assumption 1, we have $p_{j \cdot il} = p_{j \cdot l}$ in (4) and $p_{j \cdot i} = p_j$ in (7) for all $j \in \{g, g'\}$, $i \in \{1, \ldots, n\}$ and $l \in \{1, \ldots, k\}$. Therefore, it is straightforward to show that problem (1) can be equivalently formulated[11] as:

$$H_{0m} : v_m = 0 \text{ versus } H_{1m} : v_m \leqslant 0 \ \wedge \ H_{2m} : v_m \nleqslant 0 \text{ and } v_m \ngeqslant 0 \tag{14}$$

---

[11] See Anderson (1996) for a similar formulation.

for any $m \in \mathbb{N}$, where $v_m = \mathbf{T}^m(p_g - p_{g'})$ and $\mathbf{T}$ is given in (12). Since v_m is a $k \times 1$ scaled vector of contrasts, testing H_{0m} in (14) involves k multiple comparison procedures and there is a risk of size control related to a simultaneous testing of the significance of *pairwise contrasts*. To avoid size distortions, Richmond (1982) proposes to use the studentized maximum modulus (SMM) type statistic whose distribution is tabulated by Stoline and Ury (1979), and the statistic is employed by Beach and Richmond (1985) to construct confidence regions for Lorenz curve ordinates. In this paper, we combine the studentized maximum modulus statistic with an adjusted version of Pearson's (1900) statistic for assessing problem (14). Anderson (1996) employed a similar method in the context of income distributions but his methodology relies on the assumption that $(d_{g \cdot i})_{i=1}^n$ and $(d_{g' \cdot i})_{i=1}^n$ are independent, while ours is free of such a restriction.

To be more specific, suppose that Assumption 1 is satisfied. Hence, we have $\hat{p}_g \xrightarrow{p} p_g$ and $\hat{p}_{g'} \xrightarrow{p} p_{g'}$, so that the estimated contrast $\hat{v}_m = \mathbf{T}^m(\hat{p}_g - \hat{p}_{g'}) \xrightarrow{p} v_m = \mathbf{T}^m(p_g - p_{g'})$. If further H_{0m} holds, $v_m = 0$ and $\hat{v}_m$ will be close to zero for a large enough sample size. However, under H_{1m} or H_{2m}, neither v_m nor $\hat{v}_m$ will be close to zero. Therefore, one can detect whether H_{0m} is violated by looking at how far the estimated contrast $\hat{v}_m$ is from zero. Since the estimated contrast $\hat{v}_m$ will not be exactly zero under H_{0m} due to sampling error, a conventional way to proceed is to construct the test statistic from the distribution of $\hat{v}_m$. This approach is extensively discussed in Hausman (1978) and widely used in econometrics, especially in specification testing. Before we move on to the derivation of the statistics for H_{0m}, it is useful to establish the following convergence property for the estimated contrast of probabilities $\hat{p}_g - \hat{p}_{g'}$, as well as its scaled variant $\hat{v}_m = \mathbf{T}^m(\hat{p}_g - \hat{p}_{g'})$.

Lemma 1. *Suppose that Assumption* 1 *holds. For any admissible partition* $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$, *we have*:

$$\sqrt{n}[(\hat{p}_g - \hat{p}_{g'}) - (p_g - p_{g'})] \xrightarrow{d} N\left(0, \ \Sigma_g + \Sigma_{g'} - \Sigma_{gg'} - \Sigma'_{gg'}\right), \tag{15}$$

$$\text{and } \sqrt{n}(\hat{v}_m - v_m) \xrightarrow{d} N(0, \ \Omega_m), \tag{16}$$

where $\Sigma_j = \underset{n \to \infty}{p \lim}(\hat{\Sigma}_j)$, $j \in \{g, g'\}$, $\Sigma_{gg'} = \underset{n \to \infty}{p \lim}(\hat{\Sigma}_{gg'})$, $\Omega_m = \mathbf{T}^m(\Sigma_g + \Sigma_{g'} - \Sigma_{gg'} - \Sigma'_{gg'})\mathbf{T}'^m$, $\hat{\Sigma}_j$ *and* $\hat{\Sigma}_{gg'}$ *are defined in* (9) - (11).

Lemma 1 follows by the multivariate central limit theorem (MVCLT) property and the proof is presented in the appendix. It states that the estimated contrast $(\hat{p}_g - \hat{p}_{g'})$ and its scaled variant $\hat{v}_m$ are *root-n* consistent and asymptotically normal. Anderson (1996) assumes that $\Sigma_{gg'} = 0$, so we have $\Omega_m = \mathbf{T}^m(\Sigma_g + \Sigma_{g'})\mathbf{T}'^m$ in his setup. In the context of correlated samples ($\Sigma_{gg'} \neq 0$), a correction to Anderson's (1996) statistics is necessary to avoid size distortions, and the term $-(\Sigma_{gg'} + \Sigma'_{gg'})$ on the rhs of (16) is the adjustment needed.[12] In the appendix (see Lemma 4), we show that Ω_m can be consistently estimated by $\hat{\Omega}_m = \mathbf{T}^m(\hat{\Sigma}_g + \hat{\Sigma}_{g'} - \hat{\Sigma}_{gg'} - \hat{\Sigma}'_{gg'})\mathbf{T}'^m$, where $\hat{\Sigma}_j, j \in \{g, g'\}$ and $\hat{\Sigma}_{gg'}$ are given in (9)-(11). Observe that $\hat{\Sigma}_{gg'}$ is built from the contingency table obtained from the partition $\mathbf{P}_n^{(k)}(\mathbf{I})$ (thus from the distribution of the joint sample), while $\hat{\Sigma}_j$ only exploits the information from the marginal distribution of the sample of network $j \in \{g, g'\}$. The main conclusion here is that even though the cdfs (hence the pdfs) of the two networks are equal under H_{0m}, constructing the Pearson- or SMM-type statistics solely based on them, as is usually done in the literature on inequality and poverty measures, is not always the best way to go because it does not account for the correlation structure between networks.

We now focus on the derivation of the test statistics for H_{0m}.

4.1 Test Statistics and Decision Rule

Following Anderson (1996), we consider two statistics based on the estimated vector of contrasts $\hat{v}_m$ for assessing H_{0m}:

$$
\begin{aligned}
\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) &= n\hat{v}'_m\hat{\Omega}_m^-\hat{v}_m = n\hat{v}'_m\hat{P}_{k-1}\hat{D}_{k-1}^{-1}\hat{P}'_{k-1}\hat{v}_m, \\
\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) &= \max_{1 \leqslant l \leqslant k-1}\left(|\sqrt{n}\hat{Z}_{ml}|\right),
\end{aligned}
\tag{17}
$$

where $\hat{Z}_{ml}$ is the lth component of $\hat{D}_{k-1}^{-1/2}\hat{P}'_{k-1}\hat{v}_m$, $\hat{D}_{k-1}$ and $\hat{P}_{k-1}$ are given in (13). $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ in (17) is a *Pearson*-type statistic expressed as a quadratic form in $\hat{v}_m$. It differs from that in Anderson (1996) not only through the correction of the covariance matrix $\hat{\Omega}_m$, but also its direct dependence on $\mathbf{P}_n^{(k)}(\mathbf{I})$. The dependence on $\mathbf{P}_n^{(k)}(\mathbf{I})$ underscores the importance of

---

[12]Our investigation through a Monte Carlo experiment shows that failing to adjust Anderson's (1996) statistics yields overly size distorted tests when the two samples $(d_{g \cdot i})_{i=1}^n$ and $(d_{g' \cdot i})_{i=1}^n$ are correlated. In order to shorten the exposition, this exercise is omitted from this paper but it is available upon request.

controlling the size of the resulting test uniformly over the collection of admissible partitions $\mathscr{P}_A$. Uniformity over $\mathscr{P}_A$ is crucial for the asymptotic results to give a good approximation of the empirical size of the tests to the nominal level. $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ is a generalization of the SMM statistic in Stoline and Ury (1979). Besides its dependence on $\mathbf{P}_n^{(k)}(\mathbf{I})$, the expression of $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ in (17) is conceptually different from those in Stoline and Ury (1979), Beach and Richmond (1985), and Anderson (1996). For example, Beach and Richmond (1985) and Anderson (1996) defined these statistics as $\max_{1 \leqslant l \leqslant k} \left(\hat{\omega}_{m \cdot ll}^{-1/2} |\sqrt{n} \hat{v}_{ml}| \right)$, where $\hat{v}_{ml}$ is the lth component of $\hat{v}_m$ and $\hat{\omega}_{m \cdot ll}$ is the $(l,l)^{th}$ element of $\hat{\Omega}_m$. Since $\hat{v}_{ml}$ and $\hat{\omega}_{m \cdot ll}$ are not independent by construction,[13] $\max_{1 \leqslant l \leqslant k} \left(\hat{\omega}_{m \cdot ll}^{-1/2} |\sqrt{n} \hat{v}_{ml}| \right)$ does not follow a SMM distribution under H_{0m}. By contrast, the expression of $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ in (17) converges to a SMM distribution with parameter $k - 1$ and infinite degrees of freedom under H_{0m} and Assumption 1 (see Lemma 2). This is because we have adjusted this statistic as the maximum of the absolute values of $k - 1$ non-redundant linear combinations of the components of $\sqrt{n} \hat{v}_m$, where the weights are the elements of the $(k - 1) \times k$ matrix $\hat{D}_{k-1}^{-1/2} \hat{P}'_{k-1}$, while $\max_{1 \leqslant l \leqslant k} \left(\hat{\omega}_{m \cdot ll}^{-1/2} |\sqrt{n} \hat{v}_{ml}| \right)$ is obtained as the maximum of the absolute value of the k component of the scaled vector $[\mathbf{diag}(\hat{P}_{k-1} \hat{D}_{k-1}^{-1} \hat{P}'_{k-1})]^{1/2} \sqrt{n} \hat{v}_m$. Moreover, one of the fundamental differences between the two statistics in (17) is that $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ does not depend on either $\mathbf{T}$ or m (order of dominance tested),[14] while $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ depends on both.

Since $\hat{v}_m \xrightarrow{p} v_m$ under Assumption 1, it is clear from (14) that $F_g >_m F_{g'}$ if all components of $\hat{v}_m$ are less or equal to zero, with a strict inequality at least for one. Hence, the statistic $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$, which is a quadratic form in $\hat{v}_m$, if not combined with $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$, tests the equality between the cumulative distributions F_g and $F_{g'}$ and a rejection does not necessary entail *stochastic dominance*. Meanwhile, a rejection using the statistic $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ implies *stochastic dominance*. Furthermore, the test with $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ asymptotically controls the "familywise" rate of type I error in multiple comparison procedures (e.g., Richmond (1982)

---

[13] The upper α-points of the distribution of the SMM statistic, $\max_{1 \leqslant l \leqslant k} \left(\hat{\omega}_{m \cdot ll}^{-1/2} |\sqrt{n} \hat{v}_{ml}| \right)$, in Stoline and Ury (1979, Tables 1-3) are provided under the assumption that $\hat{v}_{ml}$ is independent of $\hat{\omega}_{m \cdot ll}$. However, the partitioning into classes does not preserve this independence assumption.

[14] As $\mathbf{T}$ is invertible, $\mathbf{T}^m$ is also invertible for all $m \in \mathbb{N}$ so that $\hat{v}'_m \hat{\Omega}_m^- \hat{v}_m = \hat{v}' \mathbf{T}^{m'} \mathbf{T}^{-m'} \hat{\Omega}^- \mathbf{T}^{-m} \mathbf{T}^m \hat{v} = \hat{v}' \hat{\Omega}^- \hat{v}$, i.e., $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ does not depend on either $\mathbf{T}$ nor m.

and Beach and Richmond (1985)). A combination of the two statistics informs us on whether *'stochastic dominance'* holds or not, once equality between the two distributions is rejected. Formally, as long as the two statistics are combined, one of the following three levels of decision can be reached given any admissible partition $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$:

1. if $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant c_k(\alpha)$, retain H_{0m};

2. if $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > c_k(\alpha)$ and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > s_k(\alpha)$, retain H_{1m};

3. if $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > c_k(\alpha)$ and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant s_k(\alpha)$, retain H_{2m},

where for some $\alpha \in (0, 1)$, the cut-off points $c_k(\alpha)$ and $s_k(\alpha)$ are determined such that $\mathbb{P}[\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > c_k(\alpha)] \to \alpha$ and $\mathbb{P}[\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > s_k(\alpha)] \to \alpha$ under H_{0m}, as $n \to \infty$ (at least). Tests based on the two statistics are not equally powerful against both alternatives H_{1m} and H_{2m}, especially in small samples. Indeed, in the case where one cumulative distribution is completely below the other, both tests have good power. However, if the cumulative distributions cross, the test with $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ is more powerful than those with $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$. This is because $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ is a quadratic form in $\sqrt{n}\widehat{Z}_m = \widehat{D}_{k-1}^{-1/2}\widehat{P}_{k-1}'\sqrt{n}\hat{v}_m$ while $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ is the absolute value of the maximal component of $\sqrt{n}\widehat{Z}_m \in \mathbb{R}^{k-1}$. Furthermore, from the functional forms of $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ in (17), a non-rejection by the test with $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ entails a non-rejection of those with $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$, as long as the tests are performed at the same nominal level. Thus, retaining H_{0m} when the test with $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ fails to reject it asymptotically controls the "familywise" rate of type I error. Hence, Bonferroni-type size correction for multiple comparison hypotheses is not warranted in large samples. To enhance the small-sample performance of the test, we propose a bootstrap method that is easy to implement from the observed data (see Section 4.3). But before we move on to the bootstrap results, it is informative to study the asymptotic properties of the standard tests first.

4.2 Asymptotic Properties of the tests

In this section, we characterize the large-sample properties (size and power) of the above tests of stochastic dominance. To do this, we first study the asymptotic behavior of $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$

and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ under both the null hypothesis (H_{0m}) and the alternative hypotheses $(H_{1m}$ and $H_{2m})$. Lemma 2 presents the results.

Lemma 2. *Let* $\mathbf{P}_n^{(k)}(\mathbf{I})$ *be any admissible partition in* $\mathscr{P}_A$. *Under Assumption 1, the following convergence results hold as n goes to infinity:*

(a) *if H_{0m} is satisfied, we have*

$$\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \xrightarrow{d} \chi^2(k-1), \quad \mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \xrightarrow{d} \max_{1 \leqslant l \leqslant k-1} |\mathscr{Z}_l| \sim SMM(k-1, \infty),$$

(b) *if H_{1m} or H_{2m} is satisfied, we have*

$$\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \xrightarrow{p} +\infty, \quad \mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \xrightarrow{p} +\infty,$$

where $\mathscr{Z}_l \overset{i.i.d.}{\sim} N(0,1)$ *for all $l = 1, 2, \ldots k-1$ and $SMM(k-1, \infty)$ is the studentized maximum modulus distribution with parameter $k-1$ and infinite degrees of freedom.*

Lemma 2-(a) shows that for any admissible partition $\mathbf{P}_n^{(k)}(\mathbf{I})$ in $\mathscr{P}_A$, the asymptotic distributions under H_{0m} of both statistics are nuisance parameters free. The statistic $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ has the standard χ^2 asymptotic distribution, while that of $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ is non-standard but its critical values are tabulated in Stoline and Ury (1979). Lemma 2-(b) indicates that the statistics diverge under H_{1m} or H_{2m} for any admissible partition $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$. We can now establish the following results on the uniform control of the size over $\mathscr{P}_A$ as well as test consistency for any partition $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$.

Theorem 1. *Suppose that Assumption 1 is satisfied and let $\alpha \in (0, 1)$. As the sample size n goes to infinity, the following convergence results holds:*

(a) *if H_{0m} is satisfied, then we have*

$$\limsup_{n\to\infty} \sup_{\mathscr{P}_A} \mathbb{P}[\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > \chi^2_{k-1}(\alpha)] = \alpha, \ \limsup_{n\to\infty} \sup_{\mathscr{P}_A} \mathbb{P}[\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > z_{k-1}(\alpha)] = \alpha;$$

(b) *if H_{1m} or H_{2m} is satisfied, then we have*

$$\lim_{n\to\infty} \mathbb{P}[\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > \chi^2_{k-1}(\alpha)] = 1, \quad \lim_{n\to\infty} \mathbb{P}[\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > z_{k-1}(\alpha)] = 1 \ \forall \ \mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A,$$

where $\chi^2_{k-1}(\alpha)$ and $z_{k-1}(\alpha)$ are the $(1-\alpha)^{th}$ quantiles of a $\chi^2(k-1)$-distributed and a $SMM(k-1,\infty)$-distributed random variables, respectively.

Theorem 1-(a) shows that tests based on both $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ have correct size uniformly over $\mathscr{P}_A$. Therefore, the asymptotic χ^2 and SMM critical values provide good approximations of the empirical critical values of $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ if n is large. Theorem 1-(b) indicates that both tests are consistent under H_{1m} or H_{2m} for any admissible partition $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$. However, the finite-sample size and power of the tests depend on the choice of $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$, and may not be as good as their asymptotic properties. To address this issue, we propose a bootstrap method to enhance the finite-sample properties of the tests. Section 4.3 presents the details.

4.3 Bootstrap Tests

In this section, we study the validity of the bootstrap for the statistics $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$. The usual intuition for the bootstrap requires that the empirical distribution, from which the bootstrap sample is drawn, be close to the distribution of the data under the null hypothesis. In our context, the empirical distribution used in the bootstrap sampling is the empirical distribution of the joint sample $\mathscr{D}_n = \{(d_{g \cdot i}, d_{g' \cdot i})\}_{i=1}^n$. To be more specific, the bootstrap pseudo-samples and statistics, as well as the decision rule are obtained following the above steps.

1. From the observed joint sample $\mathscr{D}_n = \{(d_{g \cdot i}, d_{g' \cdot i})\}_{i=1}^n$, obtain a partition $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$ and compute the realizations of the statistics $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$.

2. For each bootstrap sample $b = 1, \ldots, M_b$, generate the data $\mathscr{D}_n^* = \left\{ \left(d_{g \cdot i}^*, d_{g' \cdot i}^* \right) \right\}_{i=1}^n$, where $(d_{g \cdot i}^*, d_{g' \cdot i}^*)$ are drawn independently from the empirical distribution of the joint sample $\mathscr{D}_n$. From the re-sampled data and the partition $\mathbf{P}_n^{(k)}(\mathbf{I})$, compute the realizations of the bootstrap statistics $\mathbf{W}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I}))$, $\mathbf{S}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) : b = 1, \ldots, M_b :$

$$\mathbf{W}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) = n\tilde{v}_m^{*\prime}\hat{\Omega}_m^{*-}\tilde{v}_m^*, \quad \mathbf{S}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) = \max_{1 \leqslant l \leqslant k-1}\left(|\sqrt{n}\tilde{Z}_{ml}^*|\right), \quad (18)$$

where $\tilde{v}_m^* = \hat{v}_m^* - \hat{v}_m$, $\tilde{Z}_{ml}^* = \hat{Z}_{ml}^* - \hat{Z}_{ml}$; and $\hat{\Omega}_m^{*-}$, $\hat{v}_m^*$, $\hat{Z}_{ml}^*$ are the bootstrap analogues of $\hat{\Omega}_m^-$, $\hat{v}_m$, $\hat{Z}_{ml}$, respectively.

21

3. The decision rule of the bootstrap test is as follows:

    (a) if $\frac{1}{M_b}\sum_{b=1}^{M_b}\mathbb{1}\big[\mathbf{W}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) > \mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))\big] \geqslant \alpha$ where $\mathbb{1}[C] = 1$ if condition C holds and $\mathbb{1}[C] = 0$ otherwise, retain H_{0m};

    (b) if $\frac{1}{M_b}\sum_{b=1}^{M_b}\mathbb{1}\big[\mathbf{W}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) > \mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))\big] < \alpha \wedge \frac{1}{M_b}\sum_{b=1}^{M_b}\mathbb{1}\big[\mathbf{S}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) > \mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))\big] < \alpha$, retain H_{1m};

    (c) if $\frac{1}{M_b}\sum_{b=1}^{M_b}\mathbb{1}\big[\mathbf{W}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) > \mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))\big] < \alpha \wedge \frac{1}{M_b}\sum_{b=1}^{M_b}\mathbb{1}\big[\mathbf{S}_m^{*(b)}(\mathbf{P}_n^{(k)}(\mathbf{I})) > \mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))\big] \geqslant \alpha$, retain H_{2m}.

The bootstrap statistics in (18) are expressed in terms of $\tilde{v}_m^* = \hat{v}_m^* - \hat{v}_m$, rather than $\hat{v}_m^*$. This re-centering is important for the validity of the bootstrap as the expectation of $\hat{v}_m^*$ under the bootstrap data generating process is $\hat{v}_m$, which is not necessarily zero under H_{0m}. The importance of re-centering has extensively been discussed in the bootstrap literature (e.g., Hall and Horowitz (1996), Hahn (1996), Andrews (2002), Brown and Newey (2002), Inoue and Shintani (2006)).

In the remainder of the paper, the probability under the empirical distribution function of the joint sample $\mathscr{D}_n^*$ conditional on the observed data $\mathscr{D}_n$ is denoted by $\mathbb{P}^*[\cdot]$, and $\mathbb{E}^*[\cdot]$ is its corresponding expectation operator. Lemma 3 characterises the asymptotic behavior of the bootstrap statistics of stochastic dominance.

Lemma 3. *Let* $\mathbf{P}_n^{(k)}(\mathbf{I})$ *be any admissible partition in* $\mathscr{P}_A$. *Under Assumption* 1, *the following convergence results hold as* n *goes to infinity:*

(a) *if* H_{0m} *is satisfied, then we have*

$$\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{d} \chi^2(k-1) \ a.s., \ \mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{d} \max_{1\leqslant l \leqslant k-1}|\mathscr{Z}_l| \sim SMM(k-1,\infty) \ a.s.,$$

(b) *if* H_{1m} *or* H_{2m} *is satisfied, then we have*

$$\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{p} +\infty \ a.s. \ \mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{p} +\infty \ a.s.,$$

where $\mathscr{Z}_l$ *and* $SMM(k-1,\infty)$ *are defined in Lemma* 2.

Lemma 3 shows that the bootstrap provides a first-order approximation of the null limiting distributions of the statistics $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$, and is further consistent under the alternative hypotheses H_{1m} and H_{2m}. These results hold irrespective of which partition $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$ is used in the computation of the statistics. We can prove the following theorem on the consistency of the bootstrap tests.

Theorem 2. *Let* $\mathbf{P}_n^{(k)}(\mathbf{I})$ *be any admissible partition in* $\mathscr{P}_A$, *and suppose that Assumption* 1 *is satisfied. Then, the following convergence results hold as* n *goes to infinity, whether* H_{0m} *holds or not:*

$$\sup_{w \in \mathbb{R}} \left| \mathbb{P}^* \big(\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant w \big) - \mathbb{P} \big(\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant w \big) \right| \;\; \to \;\; 0 \; \text{in probability } \mathbb{P},$$

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}^* \big(\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant z \big) - \mathbb{P} (\mathbf{S}_m \big(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant z \big) \right| \;\; \to \;\; 0 \; \text{in probability } \mathbb{P}.$$

We now study the finite-sample performance (size and power) of both the standard and bootstrap tests of stochastic dominance through a Monte Carlo experiment.

5   Monte Carlo Experiment

In this section, we use simulation to examine the finite-sample size and power performance of both the standard and bootstrap tests of stochastic dominance. To shorten the exposition, we only present the results for $m = 2$ in (1). So, the null hypothesis (H_{02}) tests the equality between the two networks' distributions against second-order stochastic dominance (H_{12}), or no second-order stochastic dominance (H_{22}). The data generating process (DGP) covers the most common distributions that are used in applied work to model the degrees of networks. Precisely, the two DGPs are specified as follows.

(I). $(d_{g \cdot i}, d_{g' \cdot i})', i = 1, \ldots, n$, are drawn i.i.d. across i from a *bivariate Poisson distribution* with mean $(10, \lambda)'$ and correlation ρ. In this setup, the null hypothesis that the cdfs of $\big(d_{g \cdot i} \big)_{i=1}^n$ and $\big(d_{g' \cdot i} \big)_{i=1}^n$ are equal can be expressed as $\lambda = 10$. So, $\lambda \neq 10$ describes either H_{12} or H_{22}.

(II). $(d_{g \cdot i}, d_{g' \cdot i})', i = 1, \ldots, n$, are drawn i.i.d. across i from a *bivariate Scale-free distribu-*

tion[15] with parameters $(2.5, \gamma)'$ and correlation ρ. As in design **(I)**, the cdfs of $\left(d_{g \cdot i}\right)_{i=1}^{n}$ and $\left(d_{g' \cdot i}\right)_{i=1}^{n}$ are equal for a given ρ if and only if $\gamma = 2.5$. So, the values of $\gamma \neq 2.5$ characterize a violation of the null hypothesis.

In both setups, we vary ρ (correlation between the two samples) in $\{-0.9, -0.5, 0, 0.5, 0.9\}$, but the results do not change qualitatively with alternative choices of ρ. In all cases, the joint sample is generated using the algorithm provided by Macke et al. (2009) and Bethge and Berens (2007). As noted in Figure 2, the support of the Poisson distribution with $\lambda = 10$ is in the range 1-20, while that of the Scale-free distribution with $\gamma = 2.5$ is in the range 1-9. Hence, any admissible partition may take these ranges into account. In order to shorten the exposition, we consider two partitions for each setup. In design **(I)**, the two partitions are $k = 4$ and $k = 8$, while they are $k = 3$ and $k = 4$ in design **(II)**. Specifically, $\mathbf{P}_n^{(4)}(\mathbf{I}) := \{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3, \mathbf{I}_4\} = \{\{1, \ldots, 9\}, \{10\}, \{11\}, \{12+\}\}$ and $\mathbf{P}_n^{(8)}(\mathbf{I}) := \{\mathbf{I}_1, \ldots, \mathbf{I}_8\} = \{\{1, \ldots, 7\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14+\}\}$ in design **(I)**, and in design **(II)** we have $\mathbf{P}_n^{(3)}(\mathbf{I}) := \{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3\} = \{\{1\}, \{2\}, \{3+\}\}$ and $\mathbf{P}_n^{(4)}(\mathbf{I}) := \{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3, \mathbf{I}_4\} = \{\{1\}, \{2\}, \{3\}, \{4+\}\}$. All these partitions belong to $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$, and are thus admissible.

For the purpose of clarity and readability, we separate the analysis on the size from that on the power.

5.1 Size Properties

In this section, we analyze the empirical rejection frequencies of both the standard and bootstrap tests of stochastic dominance for various sample sizes: $n \in \{100, 200, 500\}$. In each design and for each partition $\mathbf{P}_n^{(k)}(\mathbf{I})$ specified above, the statistics $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$, $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$, $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$, and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$ are constructed as outlined in Sections 3.2, 4.1 & 4.3. The nominal level for both the standard and bootstrap tests is set at $\alpha = 5\%$ and the empirical rejection frequencies are computed with $M = 10,000$ replications. The bootstrap critical values are approximated using $M_b = 199$ pseudo samples of size n. For the standard tests, we use the $(1 - \alpha)^{th}$ quantiles of a $\chi^2(k-1)$-distributed random variable for $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$

---

[15]Note that the probability density function of a random variable D that follows a Scale-free distribution is given by $P(d) = d^{-\gamma} [\zeta(\gamma)]^{-1}, d \in \mathbb{N}$, where $\zeta(\gamma) = \sum_{d=1}^{+\infty} \frac{1}{d^{\gamma}}$ denotes the Riemann zeta function.

and a $SMM(k-1, \infty)$-distributed random variable for $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$.

Table 3 presents the results of the two designs. The first column contains the partitions $\mathbf{P}_n^{(k)}(\mathbf{I})$, and the second shows both the standard and bootstrap statistics. The other columns present, for each value of network endogeneity (ρ) and sample size n, the empirical rejection frequencies of the tests at the 5% nominal level.

First, in design **(I)** (Poisson distribution), the standard tests are slightly size distorted for $n \in \{100, 200\}$. Their maximal size rejection frequencies is around 8.7% [for $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$] and 7.2% [for $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$] with the partition $\mathbf{P}_n^{(8)}(\mathbf{I})$, but they decrease with the partition $\mathbf{P}_n^{(4)}(\mathbf{I})$ (around 6.5% and 6.2% respectively). Meanwhile, their bootstrap counterparts have rejections close to the 5% nominal level in most cases for both partitions, even with $n = 100$. However, the bootstrap tests tend to under reject when $n = 100$ and $\rho = 0.9$, but this phenomenon disappears as the sample size increases. On top of its overall good performance in small samples, our results also suggest that the bootstrap tests are less sensitive to partitioning into classes than the standard tests. Also, our results are consistent across all values of networks' endogeneity ρ.

Second, in design **(II)** (Scale-free distribution), both the standard and bootstrap tests perform quite well irrespective of the partition used and network endogeneity ρ. However, the bootstrap tests tend to be conservative when $\rho = 0.9$ and $n \in \{100, 200\}$ while the empirical rejection frequencies of the standard tests are consistently around the 5% nominal level for all sample sizes. Again, the under-rejections of the bootstrap tests observed when $\rho = 0.9$ and $n \in \{100, 200\}$ disappear as the sample size increases, as shown in the column $\rho = 0.9$ and $n = 500$ in the bottom part of the table.

5.2   Power Properties

We now study the empirical rejections of the various tests under the alternative hypothesis (power). For simplicity, we only present the power analysis for $n \in \{100, 500\}$ and $\rho \in \{0, 0.5, 0.9\}$. In design **(I)** (Poisson distribution), the power analysis is conducted in the direction of λ, where $\lambda = 10$ indicates the empirical size and $\lambda \neq 10$ indicates the empirical power. Similarly, the power analysis is conducted in the direction of γ in design **(II)** (Scale-free distribution): here $\gamma = 2.5$ indicates the empirical size, and $\gamma \neq 2.5$ characterizes the

Table 3: Empirical size of the standard and bootstrap tests at 5%

$\mathbf{P}_n^{(k)}(\mathbf{I})$		$\rho=-0.9$			$\rho=-0.5$			$\rho=0$			$\rho=0.5$			$\rho=0.9$		
	$n\to$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500

(I): _Poisson distribution_

$\mathbf{P}_n^{(4)}(\mathbf{I})$	$\mathbf{W}_m$	5.7	5.1	5.4	6.0	5.8	5.0	6.2	6.0	5.0	6.5	5.6	5.0	6.0	5.5	5.0
	$\mathbf{S}_m$	5.6	5.3	5.1	5.8	5.8	5.0	5.8	5.3	5.1	6.2	5.4	4.7	5.8	5.2	5.2
	$\mathbf{W}_m^*$	4.5	4.7	5.0	4.7	5.2	4.9	4.7	5.2	4.8	5.1	4.8	4.7	3.6	4.8	4.8
	$\mathbf{S}_m^*$	4.4	4.8	5.0	4.6	5.1	4.9	4.8	4.7	4.9	4.9	4.7	4.7	3.6	4.6	4.8
$\mathbf{P}_n^{(8)}(\mathbf{I})$	$\mathbf{W}_m$	7.8	6.6	5.9	7.8	6.5	5.6	8.1	6.5	5.9	8.7	6.6	5.7	7.4	6.1	5.3
	$\mathbf{S}_m$	6.8	6.1	5.5	7.2	5.9	5.2	7.1	6.3	5.9	7.1	5.9	5.4	6.3	5.7	5.4
	$\mathbf{W}_m^*$	4.2	4.9	5.1	4.3	4.7	4.9	4.3	4.9	5.1	4.7	4.9	5.1	2.3	4.1	4.6
	$\mathbf{S}_m^*$	4.1	5.1	5.1	4.7	4.9	4.8	4.4	5.3	5.3	4.2	4.8	5.2	2.1	4.2	5.0

(II): _Scale-free distribution_

$\mathbf{P}_n^{(k)}(\mathbf{I})$		$\rho=-0.9$			$\rho=-0.5$			$\rho=0$			$\rho=0.5$			$\rho=0.9$		
	$n\to$	100	200	500	100	200	500	100	200	500	100	200	500	100	200	500
$\mathbf{P}_n^{(3)}(\mathbf{I})$	$\mathbf{W}_m$	5.5	5.2	5.2	6.0	5.6	5.0	5.6	5.2	5.3	5.5	5.4	5.4	4.3	4.9	5.1
	$\mathbf{S}_m$	5.4	5.1	5.3	5.8	5.5	4.8	5.6	5.2	5.0	5.7	5.4	5.2	4.8	5.1	5.2
	$\mathbf{W}_m^*$	4.7	4.8	5.2	5.2	5.2	4.9	5.0	4.9	5.1	4.3	4.9	5.4	1.6	3.5	4.9
	$\mathbf{S}_m^*$	4.6	4.7	5.2	5.0	5.2	4.8	5.0	4.9	4.6	4.5	4.9	5.1	1.2	3.7	5.0
$\mathbf{P}_n^{(4)}(\mathbf{I})$	$\mathbf{W}_m$	5.8	5.2	5.4	6.0	5.7	5.2	5.9	5.4	5.2	5.4	5.4	5.1	3.9	4.4	5.0
	$\mathbf{S}_m$	5.2	5.1	5.3	5.8	5.4	4.9	5.6	5.3	4.8	5.5	5.4	5.1	4.2	4.8	5.2
	$\mathbf{W}_m^*$	4.0	4.6	5.2	4.0	5.0	4.9	3.9	4.8	4.8	2.9	4.7	4.8	1.0	2.0	4.6
	$\mathbf{S}_m^*$	3.7	4.6	5.1	4.1	4.8	4.8	3.8	4.6	4.5	2.7	4.8	5.0	0.4	2.1	4.7

empirical power at γ.

Figures 3 - 4 show the power curves of both the standard and bootstrap tests in the two partitions for design **(I)**, while Figures 5 - 6 present similar graphs for design **(II)** (Scale-free distribution). In each figure, the sub-figures (a), (c), and (e) contain the power curves of $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and its bootstrap version, while the the sub-figures (b), (d), and (f) display the power curves of $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and its bootstrap version. Each sub-figure corresponds to a value of networks' endogeneity $\rho \in \{0, 0.5, 0.9\}$.

First, when $n = 500$ and for both designs, the standard and the bootstrap tests perform similarly, irrespective of the value of ρ and the partition used (see Figure 4 and Figure 6). While the empirical power of all tests converges to 100% for large values of λ (Figure 4) and γ (Figure 6), the convergence is much lower in design **(II)** (Scale-free distribution) than in design **(I)** (Poisson distribution). This reflects the low speed of convergence in the approximation of a multinomial distribution to a multivariate normal distribution (see Lemma 1) when the original sample $\mathscr{D}_n$ is drawn from a Scale-free distribution than when it is drawn from a Poisson distribution. Although from the theory, both the standard and bootstrap tests of stochastic dominance are consistent, knowing that the empirical power of tests approaches 1 with a sample size of $n = 500$ is an interesting result.

Second, when the sample size is relatively small (here $n = 100$), substantial differences between standard and bootstrap tests appear. First, both the standard and bootstrap tests exhibit more power in design **(I)** (Poisson distribution) than in design **(II)** (Scale-free distribution). For example, for independent networks ($\rho = 0$) or low correlated networks ($\rho = 0.5$), the empirical power is low for both the standard and bootstrap tests in design **(II)** (see sub-figures (a), (b), (c) and (d) in Figure 5), while all tests exhibit more power in design **(I)** (see see sub-figures (a), (b), (c) and (d) in Figure 3). Second, within partitions, the standard and bootstrap tests perform more similarly in design **(I)** than in design **(II)**. The slightly higher power of the standard tests in Figure 3, especially for $\rho \in \{0, 0.5\}$ in partition $\mathbf{P}_n^{(8)}$, is due to their inability to control for the type-I error (see Table 3). Looking at the power of the bootstrap tests, partition $\mathbf{P}_n^{(4)}$ has a small edge over partition $\mathbf{P}_n^{(8)}$, especially for $\rho \in \{0, 0.5\}$. Mann and Wald (1942) and Williams (1950) recommended to allocate the same expected number in each cell, whilst maintaining a threshold of above 5 in order to optimize test

power. Although both partitions $\mathbf{P}_n^{(4)}$ and $\mathbf{P}_n^{(8)}$ are admissible (in the sense that a threshold of above 5 is maintained in each cell), $\mathbf{P}_n^{(4)}$ is closer to Mann and Wald's (1942) and Williams's (1950) recommendation than $\mathbf{P}_n^{(8)}$ when it comes to allocate the same expected number in each cell. Note that the power gain from using $\mathbf{P}_n^{(4)}$ over $\mathbf{P}_n^{(8)}$ decreases as: (i) ρ (networks' endogeneity) increases (see sub-figures (c)-(f) in Figure 3), or (ii) the sample size increases (see Figure 4). Finally, in design **(II)** (Scale-free distribution), while the standard tests perform similarly in the two partitions (and also outperform their bootstrap counterparts in most cases), the power of the bootstrap tests is lower with partition $\mathbf{P}_n^{(4)}$ than with $\mathbf{P}_n^{(3)}$. The power gain from using $\mathbf{P}_n^{(3)}$ over $\mathbf{P}_n^{(4)}$ can even be substantial, especially with the bootstrap test $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$ (see sub-figures (d) and (f) in Figure 5). Again, partition $\mathbf{P}_n^{(3)}$ is closer to Mann and Wald's (1942) and Williams's (1950) recommendation than partition $\mathbf{P}_n^{(4)}$.

Clearly, although bootstrapping has an overall good performance in terms of size control irrespective of which partition in $\mathscr{P}_A$ is used, our Monte Carlo results suggest that using the partition that is closer to equalizing the expected number in cells can results in a substantial power gain. Therefore, our recommendation is to follow this rule upon adjusting for the form of the distribution of the degrees, as discussed in (2)-(6) of Section 3.2.
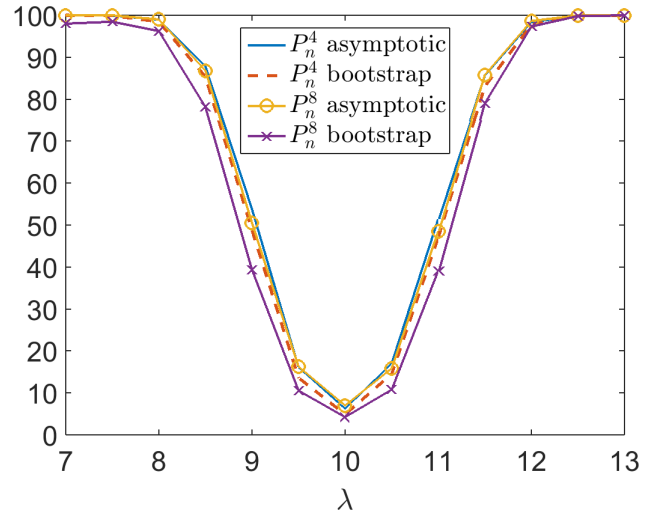
Figure 3: Power with Poisson distribution: $n = 100$
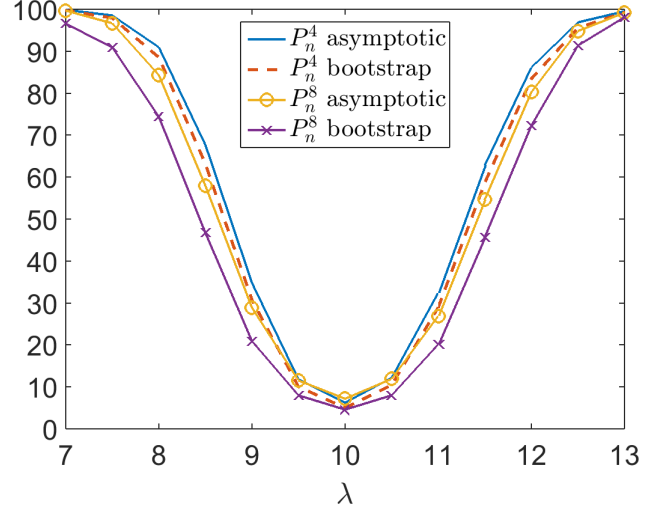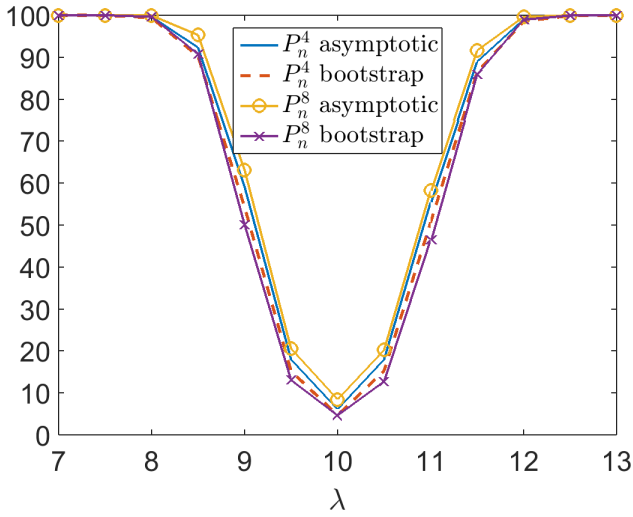
(a) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})): \rho = 0$
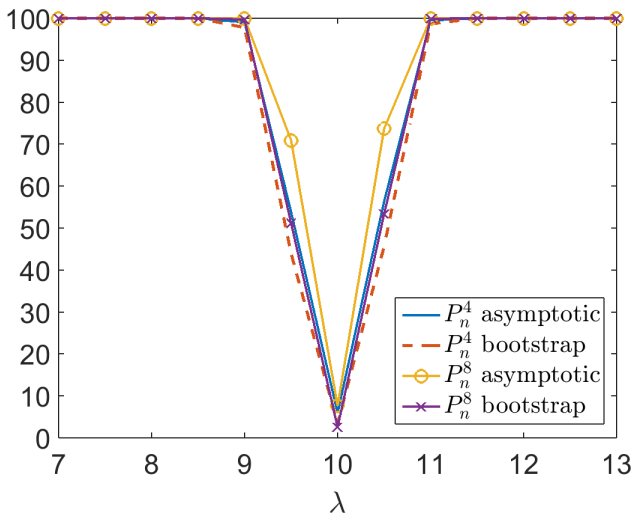
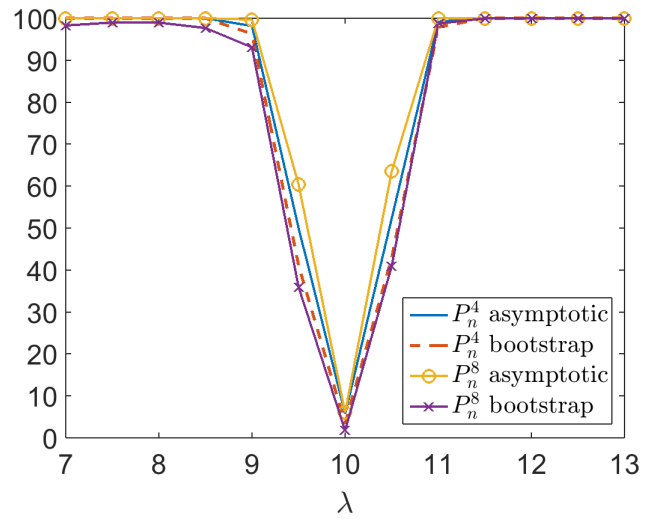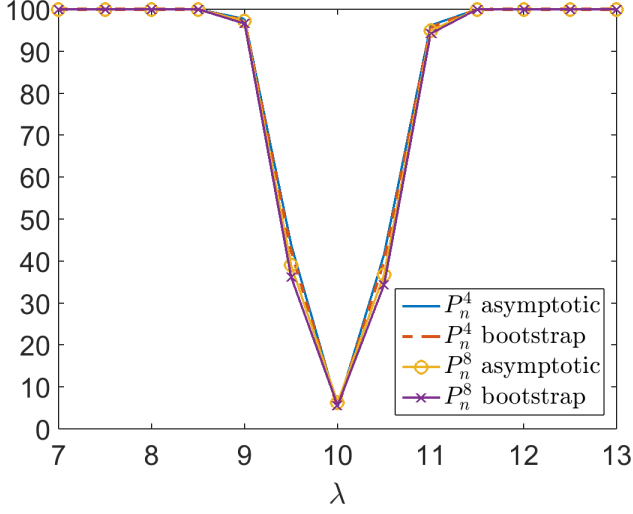(b) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})): \rho = 0$

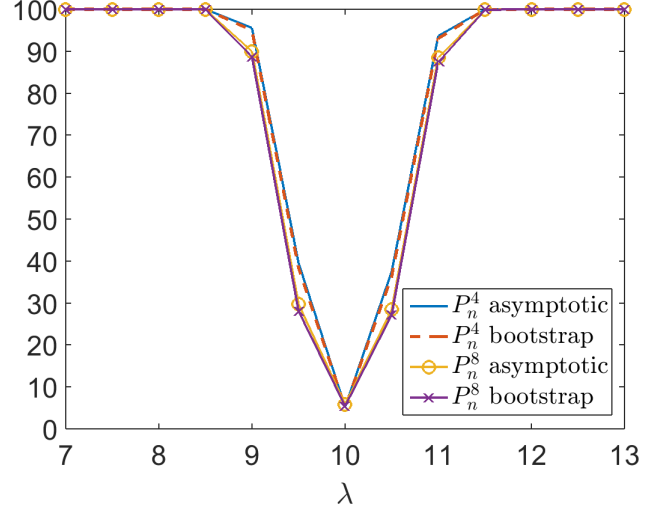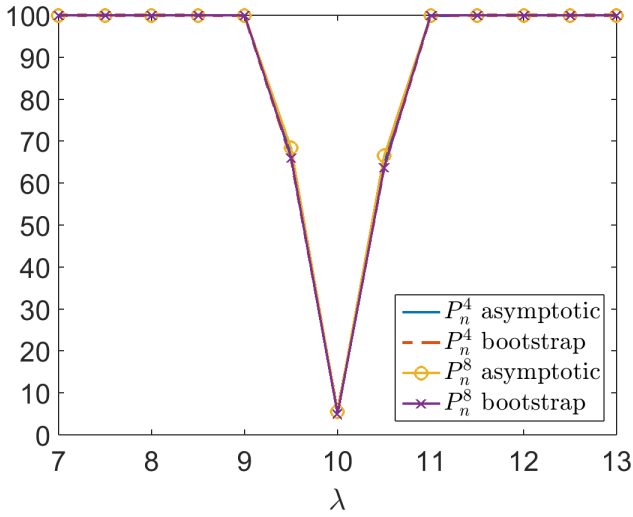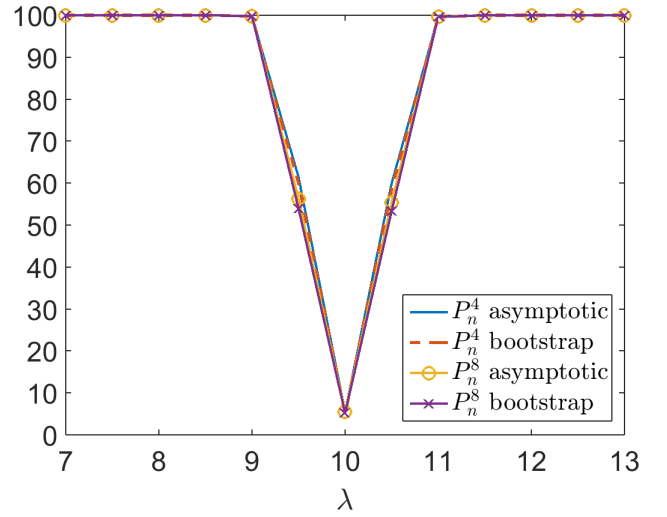(c) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})): \rho = 0.5$

(d) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})): \rho = 0.5$

(e) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})): \rho = 0.9$

(f) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})): \rho = 0.9$

Figure 4: Power with Poisson distribution: $n = 500$

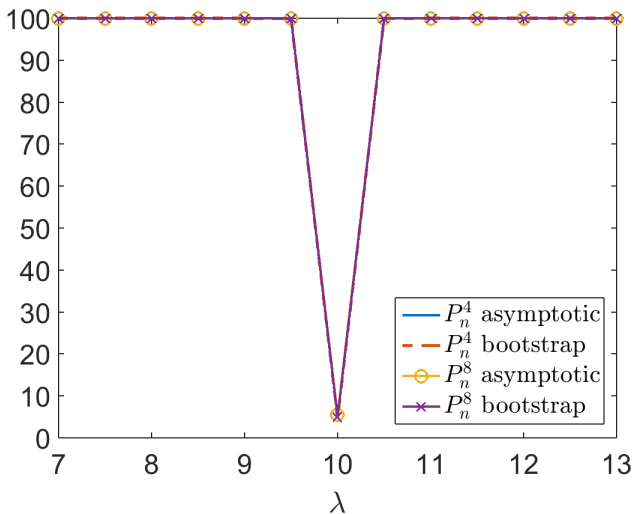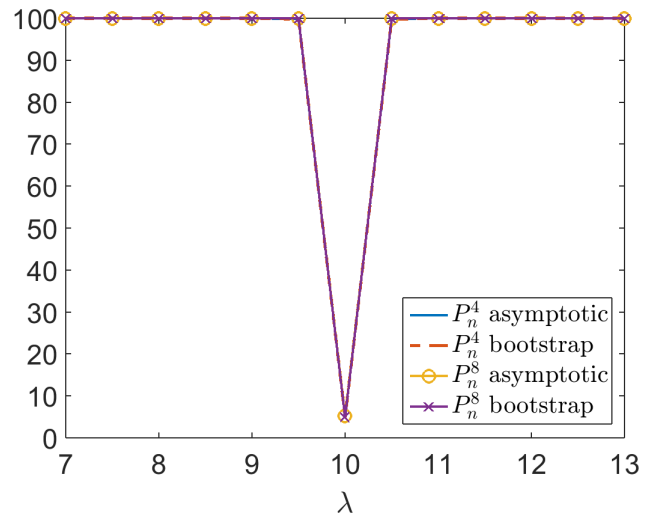(a) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0$

(b) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0$

(c) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.5$

(d) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.5$
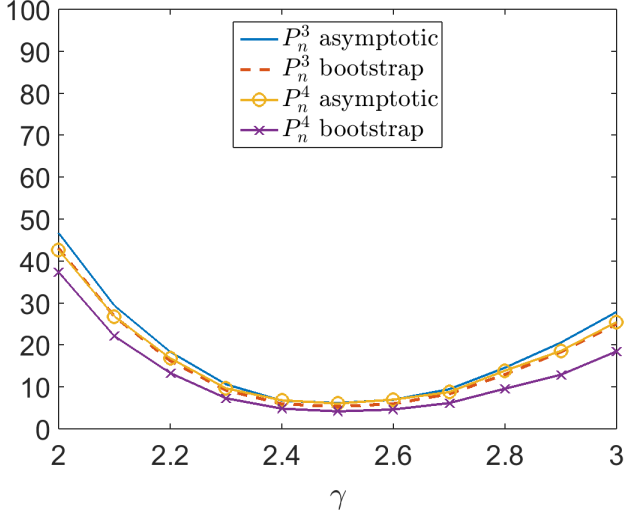
(e) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.9$
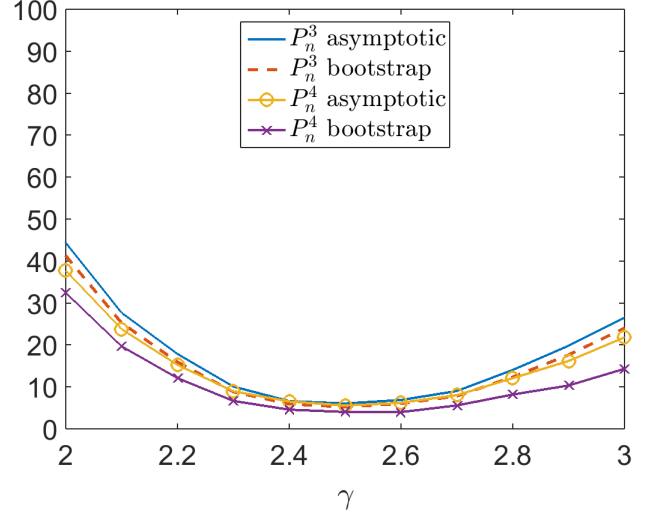
(f) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.9$
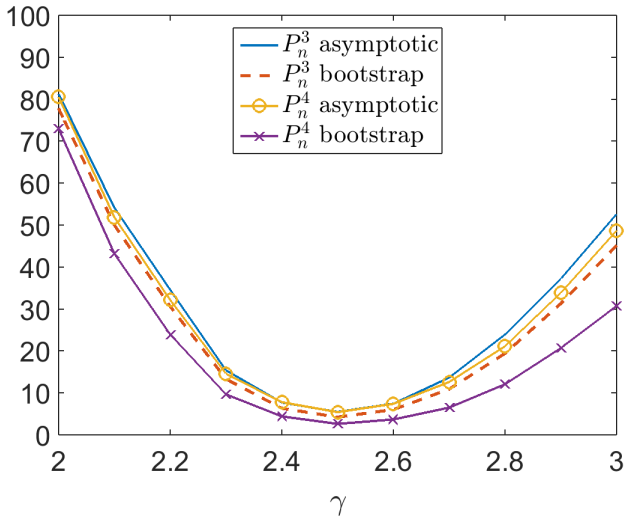
Figure 5: Power with Scale-free distribution: $n = 100$

(a) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0$

(b) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0$

(c) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.5$

(d) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.5$

(e) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.9$

(f) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.9$

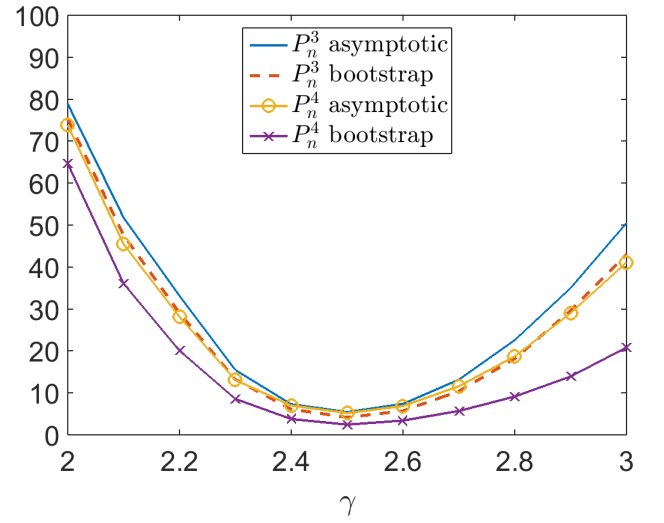Figure 6: Power with Scale-free distribution: $n = 500$

(a) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0$
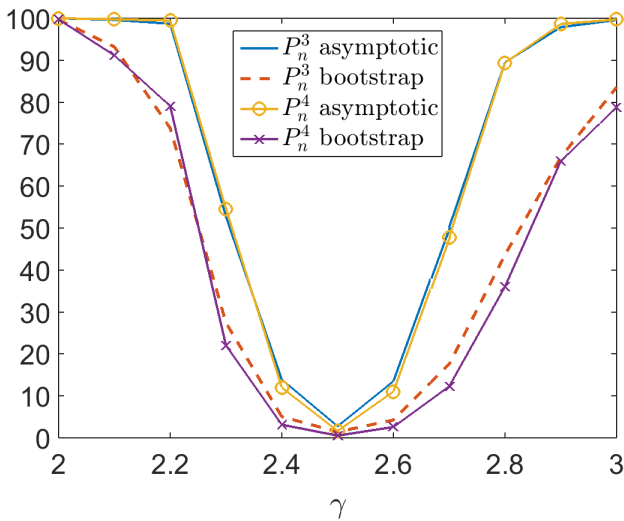
(b) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0$
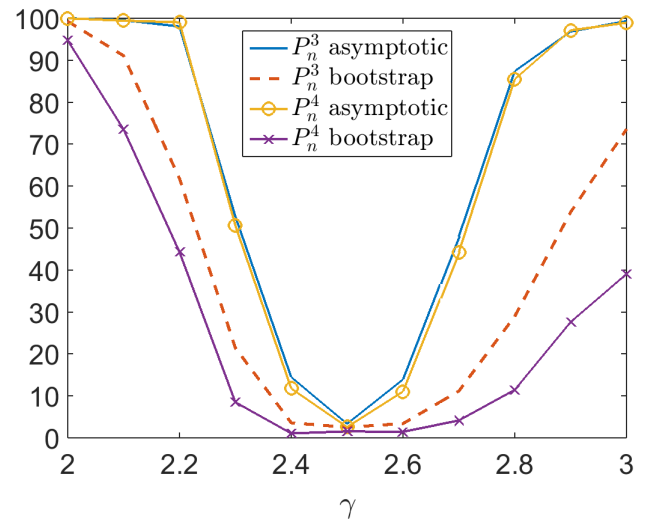
(c) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.5$

(d) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.5$
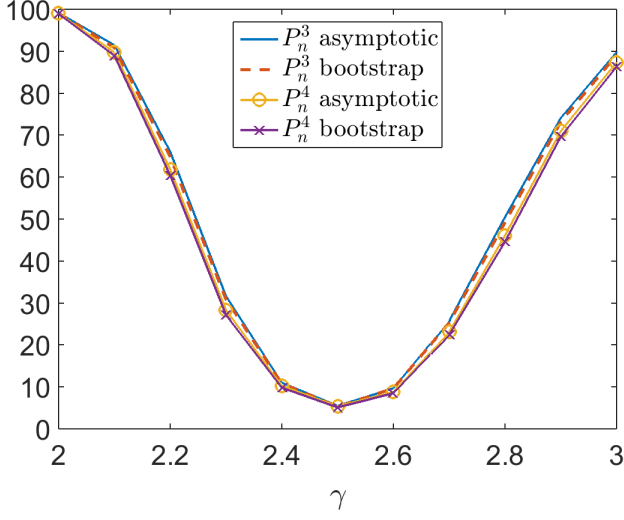
(e) $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.9$
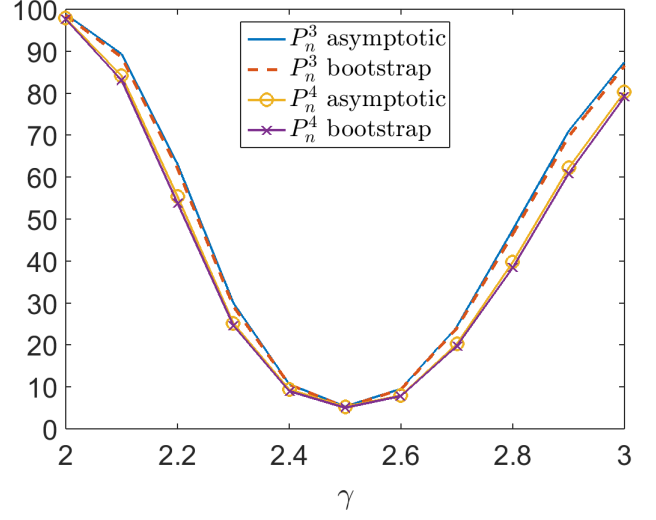
(f) $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) : \rho = 0.9$
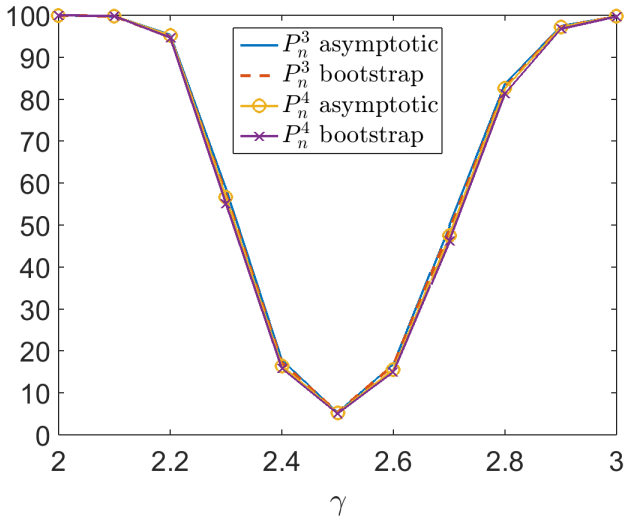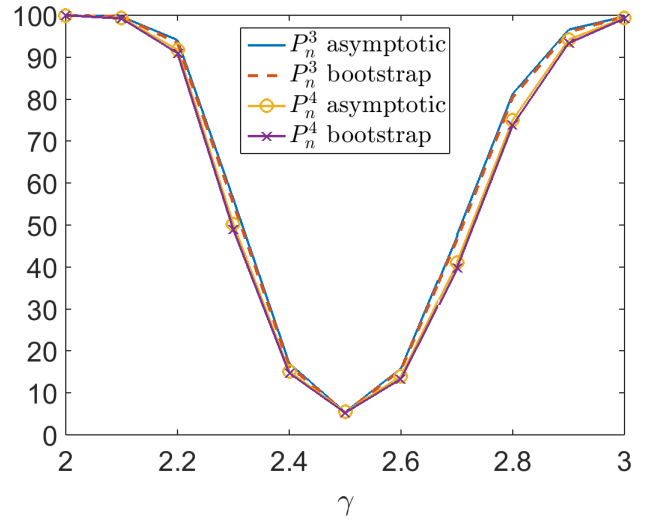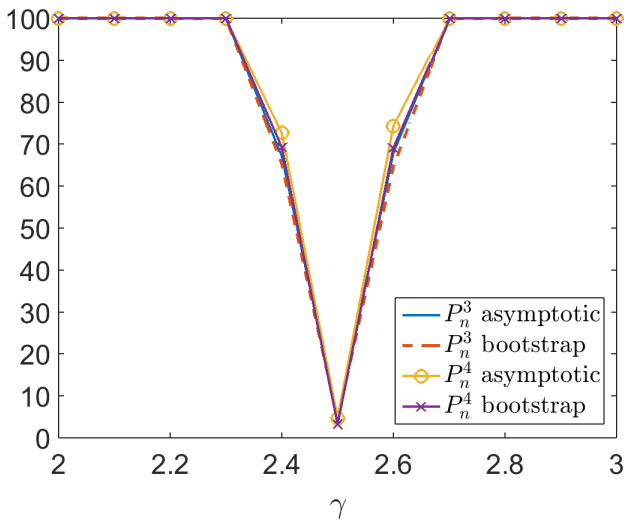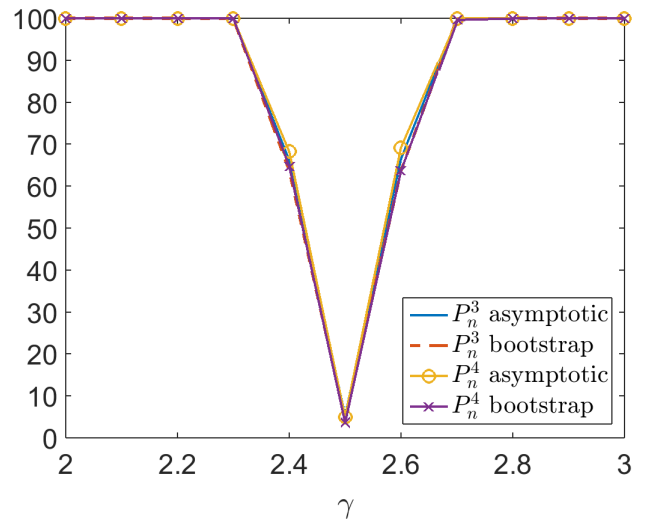
6 Empirical Illustration

Rosenzweig and Stark (1989) illustrate the strategic role that women play in smoothing consumption between villages whose income shocks are negatively correlated. In this application, we investigate whether such a role exists for sharing risk between households in rural India. In particular, we focus on testing gender differences across risk sharing networks by using the stochastic dominance criteria. Bramoulle and Kranton (2007) characterize the conditions that insure the existence of an aggregate strictly increasing (and even concave) social welfare function in risk sharing networks, meaning that these networks could be ranked in terms of *social efficiency* by applying the stochastic dominance criteria in Definition 1.

We use the data set from Banerjee et al. (2012, 2013) and Jackson et al. (2012) that comprise a random sample of households from 75 different villages in southern India. We pool the sub-samples from these villages to obtain one sample. The underlying assumption here is that the 75 sub-samples are independent across villages, but not at the household level. Each village contains on average 223 households with approximately half being sampled. Each member of a surveyed household was asked to identify members of the village with whom they engaged in a particular relationship, such as whose home they visit or with whom they go to temple. Additionally, a census on the socioeconomic characteristics– such as age, gender, religion, etc– of households was used to complete the data set; see Banerjee et al. (2012, 2013) and Jackson et al. (2012) for a detailed description of the data.

To identify risk sharing behavior we use data on the following questions: Who would come to you if he (or she) needed to borrow kerosene or rice? Who do you trust enough that if he (or she) needed to borrow 50 rupees for a day you would lend it to him (or her)? We construct female and male networks for each of the *goods lending* and *money lending* relationships as follows. We remove from the sample any person who does not name at least one connection, as it is difficult to distinguish non-response from having zero connections. We also remove any person under the age of 18. Of the remaining observations, we omit any household which does not contain at least one man and one woman. The networks are then constructed with a node representing each household. In the *female money lending network*, there is a directed link from household i to household i' if any woman in household

i has reported that she would lend money to any member (male or female) of household i', and similarly for the *male money lending network*. This means that the male and female networks have the same set of households as *nodes* and the gender corresponding to the network determines the set of *directed links*. The goods lending networks are constructed similarly. As an illustration, Figure 7 shows these networks within the households of village 1 in the data.

Figure 7: Risk Sharing Networks for Village 1

(a) Female Goods Lending

(b) Male Goods Lending

(c) Female Money lending

(d) Male Money Lending
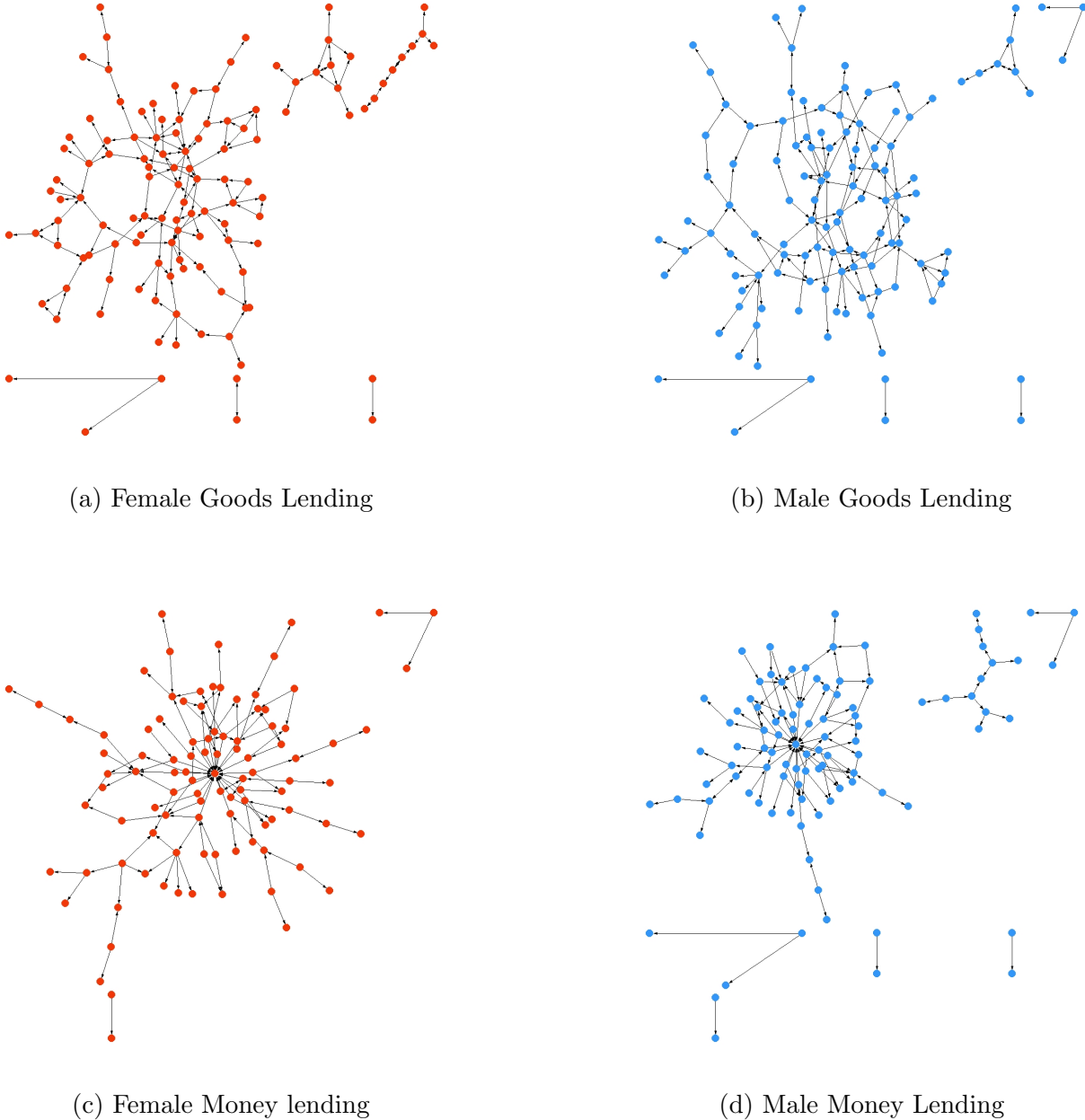
As outlined above, we conduct the tests using the pooled sample of all 75 villages. The pooled sample has size $n = 5924$ households in *goods lending networks*, and $n = 5656$ households in *money lending networks*. Table 4 summarizes the out-degree distributions of these networks as well as the correlations between *male and female networks* for both *goods lending* and *money lending*. As seen, the correlation between *male and female networks* is not small: 0.55 (for goods lending) and 0.46 (for money lending). Furthermore, in each case (goods lending and money lending) the degree distributions of both male and female networks are closer to the degree distribution of a Poisson random graph than that of a scale-free network (see Figure 2). From Sections 3.2-5, we use the following admissible partition with $k = 5$ based on Table 4:

$$\mathbf{P}_n^{(5)}(\mathbf{I}) = \{\mathbf{I}_l\}_{l=1}^5 \,, \ \mathbf{I}_l = \{l\} \text{ for } l = 1, \ldots, 4 \text{ and } \mathbf{I}_5 = \{5+\}. \tag{19}$$

In both the goods lending and money lending networks, we test whether the *female network* first- and second-order stochastically dominates the *male network*. The tests are run at the 1% and 5% nominal levels, and the bootstrap statistics critical values are approximated using $B = 199$ pseudo-samples. The results are displayed in table 5. For goods lending, both the standard and bootstrap tests are in favor of the first- and second-order stochastic dominance of the *female network* at the 1% and 5% nominal levels. However, for money lending, we could only find evidence of the first- and second-order dominance of the female network at the 5% nominal level. At the 1% nominal level, neither network dominates the other using both the standard and bootstrap tests. These results suggest that women overall tend to form denser risk sharing networks than do men, especially for goods lending. One possible explanation for this might be a higher average risk aversion among women, as documented by Borghans et al. (2009).

Table 4: Empirical Degree Distributions

Degree	Goods		Money	
	Male	Female	Male	Female
1	527	426	962	1012
2	2554	2133	2653	2509
3	1801	1831	1270	1263
4	734	1014	460	564
5	172	306	164	194
6	94	136	82	69
7	32	46	39	21
8	7	19	17	14
9	0	5	3	3
10	2	6	2	5
11	0	2	1	1
12	0	0	0	1
13	0	0	2	0
14	1	0	1	0
Obs.	5924	5924	5656	5656
Correlation	0.55		0.46	

Table 5: Stochastic dominance between female and male networks

Goods lending networks

Statistics ↓ m →	$\alpha = 0.01$		$\alpha = 0.05$	
	1	2	1	2
$\mathbf{W}_m$	310.08	310.08	310.08	310.08
$\chi^2_4(\alpha)$	13.28	13.28	9.49	9.49
$c^*_{\mathbf{W}_m}(\alpha)$	11.49	12.36	9.32	10.09
$\mathbf{S}_m$	16.92	17.01	16.92	17.01
$z_4(\alpha)$	3.02	3.02	2.49	2.49
$c^*_{\mathbf{S}_m}(\alpha)$	3.17	2.93	2.64	2.60

Money lending networks

Statistics ↓ m →	$\alpha = 0.01$		$\alpha = 0.05$	
	1	2	1	2
$\mathbf{W}_m$	19.29	19.29	19.29	19.29
$\chi^2_4(\alpha)$	13.28	13.28	9.49	9.49
$c^*_{\mathbf{W}_m}(\alpha)$	15.60	16.73	11.07	8.80
$\mathbf{S}_m$	2.92	2.59	2.92	2.59
$z_4(\alpha)$	3.02	3.02	2.49	2.49
$c^*_{\mathbf{S}_m}(\alpha)$	3.05	3.38	2.59	2.47

† $\chi^2_4(\alpha)$ and $z_4(\alpha)$ are the $(1-\alpha)^{th}$ quantiles of a chi-squared distributed random variable with 4 degrees of freedom a $SMM(4, \infty)$-distributed random variable respectively.

‡ $c^*_{\mathbf{W}_m}(\alpha)$ and $c^*_{\mathbf{S}_m}(\alpha)$ are the $(1-\alpha)^{th}$ critical values of the bootstrap statistics $\mathbf{W}^*_m$ and $\mathbf{S}^*_m$ respectively.

7 Conclusion

This paper has illustrated how stochastic dominance criteria can be used to rank networks in terms of social efficiency, and developed statistical tests for assessing these criteria. The tests proposed can be seen as a generalization of the Pearson-type and the studentized maximum modulus (SMM)-type statistics usually employed for assessing stochastic dominance criteria in the literature on income distributions, poverty and inequality measures. Our statistics differ from the prior literature not only through a correction to account for the correlation between the degree distributions of networks, but also their direct dependence on partitioning into classes. We show that a combination of the modified Pearson- and SMM-type statistics into a single decision rule is necessary to inform us on whether *stochastic dominance* holds or not, once equality between the degree distributions of the networks is rejected. As these statistics often depend on the way class intervals are allocated, controlling for type-I error uniformly over the set of all *admissible class allocations*[16] is important for the asymptotic results to give a good approximation of their empirical size to the nominal level.

We provide an analysis of both the size and power properties of the tests. On level control, we establish uniform convergence of their empirical size to the nominal level when the usual asymptotic chi-square and SMM critical values are applied. On power, we show that test consistency holds no matter which *admissible* partition is used. Finally, we provide a bootstrap method that enhances the finite-sample performance of the tests. We establish uniform consistency of the bootstrap for both the proposed Pearson- and SMM-tests irrespective of whether the null hypothesis holds or not. We present a Monte Carlo experiment that confirms our theoretical findings. Using the data set of Jackson et al. (2012) and Banerjee et al. (2012, 2013), the proposed tests were illustrated through an investigation of households' *risk sharing networks* across 75 villages in rural India. Both the goods lending and money lending networks were considered, and the gender difference within each network was our main focus. Our results suggested that women within these villages overall tend to form denser *risk sharing networks* than do men, especially for goods lending.

---

[16]By *admissible class allocation* or *admissible partition*, we mean a partition in which the minimum expected number in each cell is at least 5.

A   Appendix: Proofs

In order to establish the proofs of the lemmata and theorems of the main text, it is useful to state some basic convergence of covariance matrices $\widehat{\Sigma}_j$, $j \in \{g, g'\}$, $\widehat{\Sigma}_{gg'}$, and $\widehat{\Omega}_m$ given in (9) - (12).

Lemma 4. *Suppose that Assumption* 1 *holds. For any* $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$, *we have*:

$$(i.)\ \underset{n \to \infty}{p\lim}(\widehat{\Sigma}_j)\ =\ \Sigma_j := \begin{pmatrix} p_{j \cdot 1}(1 - p_{j \cdot 1}) & -p_{j \cdot 1}p_{j \cdot 2} & \cdots & -p_{j \cdot 1}p_{j \cdot k} \\ -p_{j \cdot 2}p_{j \cdot 1} & p_{j \cdot 2}(1 - p_{j \cdot 2}) & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ -p_{j \cdot k}p_{j \cdot 1} & -p_{j \cdot k}p_{j \cdot 2} & \cdots & p_{j \cdot k}(1 - p_{j \cdot k}) \end{pmatrix}\ \forall\ j \in \{g, g'\},$$

$$(ii.)\ \underset{n \to \infty}{p\lim}(\widehat{\Sigma}_{gg'})\ =\ \Sigma_{gg'} := \begin{pmatrix} p_{gg' \cdot 11} - p_{g \cdot 1}p_{g' \cdot 1} & p_{gg' \cdot 12} - p_{g \cdot 1}p_{g' \cdot 2} & \cdots & p_{gg' \cdot 1k} - p_{g \cdot 1}p_{g' \cdot k} \\ p_{gg' \cdot 21} - p_{g \cdot 2}p_{g' \cdot 1} & p_{gg' \cdot 22} - p_{g \cdot 2}p_{g' \cdot 2} & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ p_{gg' \cdot k1} - p_{g \cdot k}p_{g' \cdot 1} & p_{gg' \cdot k2} - p_{g \cdot k}p_{g' \cdot 2} & \cdots & p_{gg' \cdot kk} - p_{g \cdot k}p_{g' \cdot k} \end{pmatrix},$$

$$(iii.)\ \underset{n \to \infty}{p\lim}(\widehat{\Omega}_m)\ =\ \Omega_m := \mathbf{T}^m(\Sigma_g + \Sigma_{g'} - \Sigma_{gg'} - \Sigma'_{gg'})\mathbf{T}'^m.$$

Proof of Lemma 4. $(i.)$ Suppose that Assumption 1 holds and let $\mathbf{P}_n^{(k)}(\mathbf{I}) = \{\mathbf{I}_l\}_{l=1}^k \in \mathscr{P}_A$. From the i.i.d. sampling, it follows that $\hat{p}_{j \cdot l} = \frac{1}{n}\sum_{i=1}^n \mathbb{1}(d_{j \cdot i} \in \mathbf{I}_l) \overset{p}{\to} \mathbb{E}(d_{j \cdot i}) = p_{j \cdot il} \equiv p_{j \cdot l}$ for all $(j, l) \in \{g, g'\} \times \{1, \ldots, k\}$. It is clear from (9) that $\widehat{\Sigma}_j \overset{p}{\to} \Sigma_j$ for all $j \in \{g, g'\}$. The proof of $(ii.)$ follows the same steps and $(iii.)$ is implied by $(i.)$ and $(ii.)$.

Proof of Lemma 1. Let $\mathbf{P}_n^{(k)}(\mathbf{I}) = \{\mathbf{I}_l\}_{l=1}^k \in \mathscr{P}_A$ and define

$$\hat{p}\ =\ [\hat{p}'_g\ :\ \hat{p}'_{g'}]',\ p = [p'_g\ :\ p'_{g'}]', \tag{20}$$

where $\hat{p}_g = [\hat{p}_{g \cdot 1}, \ldots, \hat{p}_{g \cdot k}]' : k \times 1$, $\hat{p}_{g'} = [\hat{p}_{g' \cdot 1}, \ldots, \hat{p}_{g' \cdot k}]' : k \times 1$, $p_g = [p_{A \cdot 1}, \ldots, p_{A \cdot k}]' : k \times 1$, and $p_{g'} = [p_{B \cdot 1}, \ldots, p_{B \cdot k}]' : k \times 1$, so both $\hat{p}$ and p are $2k \times 1$ vectors obtained by stacking $\hat{p}_g$ and $\hat{p}_{g'}$ together (for $\hat{p}$) and p_g and $p_{g'}$ together (for p). From (7) - (8), we have $\hat{p}_j = \frac{1}{n}\sum_{i=1}^n u_{j \cdot i}$ and for each $j \in \{g, g'\}$, $u_{j \cdot i}$, $i = 1, \ldots, n$ are i.i.d. multinomial random variables with parameter $p_j = \mathbb{E}(u_{j \cdot i})$ under Assumption 1. Therefore, by the multivariate central limit

theorem (MVCLT), we have:

$$\sqrt{n}(\hat{p} - p) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \begin{bmatrix} u_{g \cdot i} - \mathbb{E}(u_{g \cdot i}) \\ u_{g' \cdot i} - \mathbb{E}(u_{g' \cdot i}) \end{bmatrix} \xrightarrow{d} N\left(0, \; \Sigma_p\right), \tag{21}$$

where $\Sigma_p = Avar\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \begin{bmatrix} u_{g \cdot i} - \mathbb{E}(u_{g \cdot i}) \\ u_{g' \cdot i} - \mathbb{E}(u_{g' \cdot i}) \end{bmatrix}\right) = \begin{bmatrix} \Sigma_g & \Sigma_{gg'} \\ \Sigma'_{gg'} & \Sigma_{g'} \end{bmatrix}$, Σ_j and $\Sigma_{gg'}$ are the

limits in Lemma 4. Now, let I_k be the identity matrix of order k. By noting that

$$\begin{aligned} \begin{bmatrix} I_k & -I_k \end{bmatrix} \sqrt{n}(\hat{p} - p) &= \begin{bmatrix} I_k & -I_k \end{bmatrix} \sqrt{n} \begin{bmatrix} \hat{p}_g - p_g \\ \hat{p}_{g'} - p_{g'} \end{bmatrix} \\ &= \sqrt{n}[(\hat{p}_g - \hat{p}_{g'}) - (p_g - p_{g'})], \end{aligned} \tag{22}$$

it is straightforward to see that $\sqrt{n}[(\hat{p}_g - \hat{p}_{g'}) - (p_g - p_{g'})] \xrightarrow{d} N\left[0, \; \Sigma_g + \Sigma_{g'} - (\Sigma_{gg'} + \Sigma'_{gg'})\right]$ from (21). This completes the proof of Lemma 1.

Proof of Lemma 2. Suppose that Assumption 1 holds and let $\mathbf{P}_n^{(k)}(\mathbf{I}) = \{\mathbf{I}_l\}_{l=1}^{k} \in \mathscr{P}_A$.

(a) Assume first that H_{0m} holds, i.e., $p_g = p_{g'}$. We focus on the statistic $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$. The proof for $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ can easily be adapted from Stoline and Ury (1979). From Lemmas 1 and 4, along with the expression of $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))$ in (12), it is straightforward to see that $\sqrt{n}\mathbf{T}^m[(\hat{p}_g - \hat{p}_{g'}) - (p_g - p_{g'})] \overset{H_{0m}}{=} \sqrt{n}\mathbf{T}^m(\hat{p}_g - \hat{p}_{g'}) \xrightarrow{d} \psi_m \sim N(0, \; \Omega_m)$ so that we get

$$\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \xrightarrow{d} \psi'_m \Omega_m^- \psi_m, \tag{23}$$

where $\Omega_m = \Sigma_g + \Sigma_{g'} - (\Sigma_{gg'} + \Sigma'_{gg'})$, and Ω_m^- is the generalized inverse of Ω_m. As $rank(\Omega_m) = k - 1$, there exists [see Dhrymes (1978, Proposition 3.5)] a diagonal matrix D_{k-1} whose diagonal elements are the nonzero eigenvalues of Ω_m (in decreasing order of magnitude), and a $k \times (k-1)$ matrix P_{k-1} whose columns are the (orthogonal) eigenvectors corresponding to the nonzero roots of Ω_m, such that

$$\Omega_m = P_{k-1} D_{k-1} P'_{k-1} \text{ and } \Omega_m^- = P_{k-1} D_{k-1}^{-1} P'_{k-1}. \tag{24}$$

Hence, we have: $\psi'_m \Omega_m^- \psi_m = \psi'_m P_{k-1} D_{k-1}^{-1} P'_{k-1} \psi_m = \bar{\psi}'_m D_{k-1}^{-1} \bar{\psi}_m$ from the last identity in (24), where $\bar{\psi}_m = P'_{k-1} \psi_m$. Since $\psi_m \sim N(0, \; \Omega_m)$, we have $\bar{\psi}_m \sim N\left(0, \; P'_{k-1} \Omega_m P_{k-1}\right) = N\left(0, \; P'_{k-1} P_{k-1} D_{k-1} P'_{k-1} P_{k-1}\right) = N(0, \; D_{k-1})$ from the first identity in (24), where $P'_{k-1} P_{k-1} =$

I_{k-1}. Therefore, $D_{k-1}^{-1/2}\bar{\psi}_m \sim N\left(0,\ I_{k-1}\right)$ so that $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \overset{d}{\to} \bar{\psi}_m' D_{k-1}^{-1}\bar{\psi}_m \sim \chi^2(k-1)$, as stated.

(b) Assume now that H_{1m} or H_{2m} is true. Hence, we have $p_g - p_{g'} \neq 0$ so that $\hat{v}_m \overset{p}{\to} v_m = \mathbf{T}^m\left(p_g - p_{g'}\right) \neq 0$. Furthermore, as $\hat{\Omega}_m \overset{p}{\to} \Omega_m$, it is clear that $\hat{v}_m'\hat{\Omega}_m^-\hat{v}_m \overset{p}{\to} (p_g - p_g)'\mathbf{T}^{m'}\Omega_m^-\mathbf{T}^m\left(p_g - p_{g'}\right) > 0$ because $rank(\Omega_m^-) = k-1$. Therefore, we find $\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) = n\hat{v}_m'\hat{\Omega}_m^-\hat{v}_m \overset{p}{\to} +\infty$. Similarly, we can see that $\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \overset{p}{\to} +\infty$. This completes the proof of Lemma 2.

Proof of Theorem 1. (a) Suppose first that H_{0m} holds. Since $\mathscr{P}_A$ is a discrete and finite set of collection of partitions $\mathbf{P}_n^{(k)}(\mathbf{I})$, the sequence of probabilities $\alpha_{1,n}^{(k)}[\mathbf{P}_n^{(k)}(\mathbf{I}), \mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))] = \mathbb{P}[\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > \chi_{k-1}^2(\alpha)] \in [0,1]$ and $\alpha_{2,n}^{(k)}[\mathbf{P}_n^{(k)}(\mathbf{I}), \mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))] = \mathbb{P}[\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > z_{k-1}(\alpha)] \in [0,1]$ can be *ordered* for all possible collections $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$. Therefore, there are sequences $\widetilde{\mathbf{P}}_n^{(k)}, \breve{\mathbf{P}}_n^{(k)} \in \mathscr{P}_A$ and subsequences $\{\pi_n : n \geqslant 1\}, \{\breve{\pi}_n : n \geqslant 1\}$ of $\{n : n \geqslant 1\}$ such that

$$
\begin{aligned}
\limsup_{n\to\infty} \sup_{\mathscr{P}_A} \alpha_{1,n}^{(k)}[\mathbf{P}_n^{(k)}(\mathbf{I}), \mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))] \quad &:= \quad \limsup_{n\to\infty} \sup_{\mathscr{P}_A} \mathbb{P}[\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > \chi_{k-1}^2(\alpha)] \\
&= \quad \limsup_{n\to\infty} \mathbb{P}[\mathbf{W}_m(\widetilde{\mathbf{P}}_n^{(k)}) > \chi_{k-1}^2(\alpha)] \\
&= \quad \lim_{n\to\infty} \mathbb{P}[\mathbf{W}_m(\widetilde{\mathbf{P}}_{\pi_n}^{(k)}) > \chi_{k-1}^2(\alpha)]. \qquad (25) \\
\limsup_{n\to\infty} \sup_{\mathscr{P}_A} \alpha_{2,n}^{(k)}[\mathbf{P}_n^{(k)}(\mathbf{I}), \mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I}))] \quad &:= \quad \limsup_{n\to\infty} \sup_{\mathscr{P}_A} \mathbb{P}[\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > z_{k-1}(\alpha)] \\
&= \quad \limsup_{n\to\infty} \mathbb{P}[\mathbf{S}_m(\breve{\mathbf{P}}_n^{(k)}) > z_{k-1}(\alpha)] \\
&= \quad \lim_{n\to\infty} \mathbb{P}[\mathbf{S}_m(\breve{\mathbf{P}}_{\breve{\pi}_n}^{(k)}) > z_{k-1}(\alpha)]. \qquad (26)
\end{aligned}
$$

But from Lemma 2-(a), we have $\lim_{n\to\infty} \mathbb{P}[\mathbf{W}_m(\widetilde{\mathbf{P}}_{\pi_n}^{(k)}) > \chi_{k-1}^2(\alpha)] = \mathbb{P}[\chi_{k-1}^2 > \chi_{k-1}^2(\alpha)] = \alpha$ and $\lim_{n\to\infty} \mathbb{P}[\mathbf{S}_m(\breve{\mathbf{P}}_{\breve{\pi}_n}^{(k)}) > z_{k-1}(\alpha)] = \mathbb{P}[SMM(k,\infty) > z_{k-1}(\alpha)] = \alpha$. Using (25)-(26), we get:

$$\limsup_{n\to\infty} \sup_{\mathscr{P}_A} \mathbb{P}[\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > \chi_{k-1}^2(\alpha)] = \alpha \text{ and } \limsup_{n\to\infty} \sup_{\mathscr{P}_A} \mathbb{P}[\mathbf{S}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) > z_{k-1}(\alpha)] = \alpha.$$

(b) Under H_{1m} or H_{2m}, the results follow immediately from Lemma 2-(b).

Proof of Lemma 3. We prove the results for $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$. The proof for $\mathbf{S}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$ can be constructed in a similar way. First, we can write the bootstrap statistic $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$ as

$$
\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \quad = \quad n\tilde{v}_m^{*\prime}\hat{\Omega}_m^{*-}\tilde{v}_m^* = n(\hat{v}_m^* - \hat{v}_m)'\hat{\Omega}_m^{*-}(\hat{v}_m^* - \hat{v}_m). \qquad (27)
$$

(a) Suppose first that H_{0m} holds and let $S_m^* = \sqrt{n}(\hat{v}_m^* - \hat{v}_m)$. We can express S_m^* as:

$$S_m^* = \sum_{i=1}^{n} R_{m,i}^*, \text{ where } R_{m,i}^* = \frac{1}{\sqrt{n}} \mathbf{T}^m \Big[(d_{g\cdot i}^* - d_{g'\cdot i}^*) - \frac{1}{n} \sum_{i=1}^{n} (u_{g\cdot i} - u_{g'\cdot i}) \Big].$$

Moreover, from the i.i.d. sampling under $\mathbb{P}^*$, we have $\mathbb{E}^*(d_{g\cdot i}^* - d_{g'\cdot i}^*) = \frac{1}{n}\sum_{i=1}^{n}(u_{g\cdot i} - u_{g'\cdot i})$, so that $R_{m,i}^*$ can be expressed as $R_{m,i}^* = \frac{1}{\sqrt{n}}\mathbf{T}^m \big[d_{g\cdot i}^* - d_{g'\cdot i}^* - \mathbb{E}^*(d_{g\cdot i}^* - d_{g'\cdot i}^*)\big]$, i.e., $\{R_{m,i}^*\}_{i=1}^{n}$ are also i.i.d under $\mathbb{P}^*$. We want to verify the conditions of the Liapunov Central Limit Theorem for S_m^*.

(a) By definition, it is straightforward to see that $\mathbb{E}^*(R_{m,i}^*) = 0$.

(b) $\mathbb{E}^*(R_{m,i}^{*2}) = var^*(R_{m,i}^*) = n^{-1}\hat{\Omega}_m < \infty$ a.s.

(c) Finally, we need to show that $\lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}^*[\|R_{m,i}^*\|^{2+\delta}] = 0$ a.s. for some $\delta > 0$. We have:

$$
\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}^*[\|R_{m,i}^*\|^{2+\delta}] &\leq cn^{-\frac{\delta}{2}}n^{-1}\sum_{i=1}^{n} \mathbb{E}^*\Big[\|\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta} + \|\frac{1}{n}\sum_{i=1}^{n}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\|^{2+\delta}\Big] \\
&= cn^{-\frac{\delta}{2}}\mathbb{E}^*\Big[\|\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta}\Big] + cn^{-\frac{\delta}{2}}\Big\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\Big\|^{2+\delta}
\end{aligned}
$$

for a large enough constant $c \in \mathbb{R}^+$.

First, we have $\frac{1}{n}\sum_{i=1}^{n}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i}) \xrightarrow{p} \mathbf{T}^m(p_g - p_{g'}) = v_m = 0$ under Assumption 1 and H_{0m}. So, the second term of the last equality in the above equation is such that $cn^{-\frac{\delta}{2}}\big\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\big\|^{2+\delta} \xrightarrow{p} 0$ since $cn^{-\frac{\delta}{2}} \to 0$ when $n \to \infty$. For the first term, we note that $\mathbb{E}^*\big[\|\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta}\big] \xrightarrow{p^*} \big\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\big\|^{2+\delta}$ and we know that $\big\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\big\|^{2+\delta} \xrightarrow{p} \big\|\mathbf{T}^m(p_g - p_{g'})\big\|^{2+\delta} = \|v_m\|^{2+\delta} = 0$ when H_{0m} holds. So, we get $cn^{-\frac{\delta}{2}}\mathbb{E}^*\big[\|\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta}\big] \xrightarrow{p} 0$ a.s. As a result, we have $\lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}^*[\|R_{m,i}^*\|^{2+\delta}] = 0$ a.s. as required.

Since $\hat{\Omega}_m^* - \hat{\Omega}_m \mid \mathscr{D}_n \xrightarrow{a.s.} 0$, $\hat{\Omega}_m \xrightarrow{p} \Omega_m$, and the conditions of the Liapunov CLT are satisfied, we have

$$S_m^* \mid \mathscr{D}_n \xrightarrow{d} \psi_m \sim N(0, \Omega_m) \text{ a.s.}$$

Now, we want to show that $\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{d} \chi^2(k-1)$ a.s. for any $\mathbf{P}_n^{(k)}(\mathbf{I}) \in \mathscr{P}_A$. From (27) and the fact that $\hat{\Omega}_m^{*-} \mid \mathscr{D}_n \xrightarrow{p} \Omega_m$ a.s., it is straightforward to see that

$$\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \mid \mathscr{D}_n = S_m^{*'}\hat{\Omega}_m^{*-}S_m^* \mid \mathscr{D}_n \xrightarrow{d} \psi_m'\Omega_m^-\psi_m \text{ a.s.} \tag{28}$$

42

Since we have $\psi'_m \Omega^-_m \psi_m \sim \chi^2(k-1)$ by Lemma 2, it is clear that $\mathbf{W}^*_m(\mathbf{P}^{(k)}_n(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{d} \chi^2(k-1)$ a.s. for all $\mathbf{P}^{(k)}_n(\mathbf{I}) \in \mathscr{P}_A$, as stated.

(b) Suppose now that H_{0m} fails, i.e., H_{1m} or H_{2m} holds. It is easy to see from the proof in (a) that $\frac{1}{\sqrt{n}} S^{*'}_m \mid \mathscr{D}_n \xrightarrow{p} v_m$ a.s., $\hat{\Omega}^*_m \mid \mathscr{D}_n \xrightarrow{p} \Omega_m$ a.s. so that $\frac{1}{n} \mathbf{W}^*_m(\mathbf{P}^{(k)}_n(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{a.s.} v'_m \Omega^-_m v_m > 0$ because $v_m \neq 0$ under H_{1m} or H_{2m}. Therefore, we have $\mathbf{W}^*_m(\mathbf{P}^{(k)}_n(\mathbf{I})) \mid \mathscr{D}_n \xrightarrow{p} +\infty$ a.s. under H_{1m} or H_{2m} for any $\mathbf{P}^{(k)}_n(\mathbf{I}) \in \mathscr{P}_A$, as required.

Proof of Theorem 2. As in Lemma 3, we will prove the results for $\mathbf{W}^*_m(\mathbf{P}^{(k)}_n(\mathbf{I}))$. The proof for $\mathbf{S}^*_m(\mathbf{P}^{(k)}_n(\mathbf{I}))$ can be constructed in a similar way.

(a) Suppose first that H_{0m} holds. We know from Lemma 3 that $\hat{\Omega}^*_m - \hat{\Omega}_m \mid \mathscr{D}_n \xrightarrow{a.s.} 0$ and $\hat{\Omega}_m$ has rank $k-1$ by construction. Hence, $\hat{\Omega}^*_m$ also has rank $k-1$ a.s. Therefore, from Dhrymes (1978, Proposition 3.5) there exists a diagonal matrix $\hat{D}^*_{k-1}$ whose diagonal elements are the nonzero eigenvalues of $\hat{\Omega}^*_m$ (in decreasing order of magnitude), a $k \times (k-1)$ matrix $\hat{P}^*_{k-1}$ whose columns are the (orthogonal) eigenvectors corresponding to the nonzero roots of $\hat{\Omega}^*_m$, such that

$$\hat{\Omega}^*_m = \hat{P}^*_{k-1}\hat{D}^*_{k-1}\hat{P}^{*'}_{k-1} \text{ and } \hat{\Omega}^{*-}_m = \hat{P}^*_{k-1}\hat{D}^{*-1}_{k-1}\hat{P}^{*'}_{k-1}, \tag{29}$$

where $\hat{P}^*_{k-1}$ and $\hat{D}^*_{k-1}$ satisfy the following convergence:

$$\hat{P}^*_{k-1} \mid \mathscr{D}_n \xrightarrow{p} P_{k-1} \text{ a.s.. and } \hat{D}^*_{k-1} \mid \mathscr{D}_n \xrightarrow{p} D_{k-1} \text{ a.s.,} \tag{30}$$

where P_{k-1} and D_{k-1} are the matrices defined in equation (24) [in the proof of Lemma 2]. Now, from the proof of Lemma 3, we can express $\mathbf{W}^*_m(\mathbf{P}^{(k)}_n(\mathbf{I}))$ as:

$$\mathbf{W}^*_m(\mathbf{P}^{(k)}_n(\mathbf{I})) = S^{*'}_m \hat{\Omega}^{*-}_m S^*_m = \tilde{S}^{*'}_m \tilde{S}^*_m, \tag{31}$$

where $\tilde{S}^*_m = \hat{D}^{*-1/2}_{k-1} \hat{P}^{*'}_{k-1} S^*_m = \sum_{i=1}^{n} \tilde{R}^*_{m,i}$ and $\{\tilde{R}^*_{m,i}\}_{i=1}^{n}$ are also i.i.d under $\mathbb{P}^*$. By adapting the proof of the Liapunov Central Limit Theorem in Lemma 3, we have

$$\tilde{S}^*_m \mid \mathscr{D}_n \xrightarrow{d} N(0, I_{k-1}) \text{ a.s.} \tag{32}$$

Moreover, since $\{\tilde{R}^*_{m,i}\}_{i=1}^{n}$ are i.i.d under $\mathbb{P}^*$ with finite second moments, from the Berry-Esseen theorem for sums of independent random vectors, we have

$$\sup_{x \in \mathbb{R}^{k-1}} \left| \mathbb{P}^*(\tilde{S}^*_m \leqslant x) - \Phi(x) \right| \leqslant \frac{c(k)}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}^*[\|\tilde{R}^*_{m,i}\|^{2+\delta}], \tag{33}$$

43

where $c(k)$ is a constant that depends on k ($=$ dimension of $\widetilde{S}_m^*$), $\Phi(\cdot) \equiv$ cdf of $N(0, I_{k-1})$. Moreover, by adapting the proof of the Liapunov Central Limit Theorem in step (c) of the proof of Lemma 3, we have

$$\sum_{i=1}^{n} \mathbb{E}^*[\|\widetilde{R}_{m,i}^*\|^{2+\delta}] = \sum_{i=1}^{n} \mathbb{E}^*\Big[\|\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}R_{m,i}^*\|^{2+\delta}\Big] \leqslant cn^{-\frac{\delta}{2}}n^{-1}\sum_{i=1}^{n}\mathbb{E}^*\Big[\|\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta}\Big]$$

$$+\|\frac{1}{n}\sum_{i=1}^{n}\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\|^{2+\delta}\Big] = cn^{-\frac{\delta}{2}}\mathbb{E}^*\Big[\|\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta} +$$

$$cn^{-\frac{\delta}{2}}\Big\|\frac{1}{n}\sum_{i=1}^{n}\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\Big\|^{2+\delta} \tag{34}$$

for a large enough constant $c \in \mathbb{R}^+$. However, the second term of the last equality in (34) is such that:

$$cn^{-\frac{\delta}{2}}\Big\|\frac{1}{n}\sum_{i=1}^{n}\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\Big\|^{2+\delta} \xrightarrow{p} 0 \ a.s.$$

because $\Big\|\frac{1}{n}\sum_{i=1}^{n}\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\Big\|^{2+\delta} \xrightarrow{a.s.} \Big\|D_{k-1}^{-1/2}P_{k-1}'v_m\Big\|^{2+\delta} = 0$ under H_{0m} and $cn^{-\frac{\delta}{2}} \to 0$ as $n \to \infty$. Similarly, the first term of the last equality in (34) is such that:

$$cn^{-\frac{\delta}{2}}\mathbb{E}^*\Big[\|\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta}\Big] \xrightarrow{p} 0 \ a.s.$$

because $\mathbb{E}^*\Big[\|\hat{D}_{k-1}^{*\,-1/2}\hat{P}_{k-1}^{*\prime}\mathbf{T}^m(d_{g\cdot i}^* - d_{g'\cdot i}^*)\|^{2+\delta}\Big] \xrightarrow{p^*} \Big\|\frac{1}{n}\sum_{i=1}^{n}\hat{D}_{k-1}^{-1/2}\hat{P}_{k-1}'\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\Big\|^{2+\delta}$ and $\Big\|\frac{1}{n}\sum_{i=1}^{n}\hat{D}_{k-1}^{-1/2}\hat{P}_{k-1}'\mathbf{T}^m(u_{g\cdot i} - u_{g'\cdot i})\Big\|^{2+\delta} \xrightarrow{p} \Big\|D_{k-1}^{-1/2}P_{k-1}'v_m\Big\|^{2+\delta} = 0$ under H_{0m}; and in addition $cn^{-\frac{\delta}{2}} \to 0$ as $n \to \infty$. Therefore, we have $\sum_{i=1}^{n}\mathbb{E}^*[\|\widetilde{R}_{m,i}^*\|^{2+\delta}] \mid \mathscr{D}_n \xrightarrow{p} 0$ a.s. in prob-$\mathbb{P}$, which entails that $\frac{c(k)}{\sqrt{n}}\sum_{i=1}^{n}\mathbb{E}^*[\|\widetilde{R}_{m,i}^*\|^{2+\delta}] \mid \mathscr{D}_n \xrightarrow{p} 0$ a.s. in prob-$\mathbb{P}$. From (33), it is clear that we have

$$\sup_{x \in \mathbb{R}^{k-1}}\Big|\mathbb{P}^*(\widetilde{S}_m^* \leqslant x) - \Phi(x)\Big| \xrightarrow{p} 0 \text{ in prob-}\mathbb{P}. \tag{35}$$

Now, by using (31), we can write $\mathbb{P}^*(\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant w)$ as: $\mathbb{P}^*(\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant w) = \mathbb{P}^*(\widetilde{S}_m^* \in \mathscr{C}_w)$ where $\mathscr{C}_w = \{x \in \mathbb{R}^{k-1} : x'x \leqslant w\}$ are convex sets in $\mathbb{R}^{k-1}$. From Bhattacharya and Rao (1976, Corollary 3.2), we have $\sup_{w \in \mathbb{R}}\Phi\big((\partial\mathscr{C}_w)^\epsilon\big) \leqslant d.\epsilon$ for some constant d and $\epsilon > 0$. Hence, Bhattacharya and Ghosh (1978, Theorem 1) holds with $W_n \equiv \mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I}))$ and $B \equiv \mathscr{C}_w$, thus

$$\sup_{w \in \mathbb{R}}\Big|\mathbb{P}^*(\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant w) - G_{k-1}(w)\Big| \xrightarrow{p} 0 \text{ in prob-}\mathbb{P}, \tag{36}$$

44

where $G_{k-1}(\cdot) \equiv$ cdf of $\chi^2(k-1)$. Finally, we have:

$$\sup_{w \in \mathbb{R}} \left| \mathbb{P}^*(\mathbf{W}_m^*(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant w) - \mathbb{P}(\mathbf{W}_m(\mathbf{P}_n^{(k)}(\mathbf{I})) \leqslant w) \right| \xrightarrow{p} 0 \text{ in prob-}\mathbb{P} \text{ by Lemma 2.}$$

(b) Under H_{1m} or H_{2m}, the results follow straightforwardly from Lemma 2-(b) and Lemma 3-(b), so the proof is omitted.

Bibliography

Aliprantis, D. and F. G.-C. Richter (2013). Evidence of neighborhood effects from MTO: LATEs of neighborhood quality. *Review of Economics and Statistics 88*(3), 389–432.

Anderson, G. (1996). Nonparametric tests of stochastic dominance in income distributions. *Econometrica 64*(5), pp. 1183–1193.

Andrews, D. W. (2002). Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators. *Econometrica 70*(1), 119–162.

Atkinson, A. B. (1987). On the measurement of poverty. *Econometrica: Journal of the Econometric Society 55*(4), 749–764.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2012, January). The diffusion of microfinance. NBER Working Papers 17743, National Bureau of Economic Research, Inc.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(6144), 1236498: 1–7.

Barrett, G. F. and S. G. Donald (2003). Consistent tests for stochastic dominance. *Econometrica 71*(1), 71–104.

Barrett, G. F., S. G. Donald, and D. Bhattacharya (2014). Consistent nonparametric tests for lorenz dominance. *Journal of Business & Economic Statistics 32*(1), 1–13.

Beach, C. M. and J. Richmond (1985). Joint confidence intervals for income shares and lorenz curves. *International Economic Review 26*, 439–450.

Bethge, M. and P. Berens (2007). Near-maximum entropy models for binary neural representations of natural images. In *Advances in Neural Information Processing Systems*, pp. 97–104.

Bhattacharya, R. N. and J. K. Ghosh (1978). On the validity of the formal Edgeworth expansion. *The Annals of Statistics 6*(2), 434–451.

Bhattacharya, R. N. and R. R. Rao (1976). *Normal approximation and asymptotic expansions.* John Wiley & Sons, Inc.

Bickel, P. J. and A. Chen (2009). A nonparametric view of network models and newman–girvan and other modularities. *Proceedings of the National Academy of Sciences 106*(50), 21068–21073.

Bickel, P. J., A. Chen, E. Levina, et al. (2011). The method of moments and degree distributions for network models. *The Annals of Statistics 39*(5), 2280–2301.

Blume, L. E., W. A. Brock, S. N. Durlauf, and R. Jayaraman (2015). Linear social interactions models. *Journal of Political Economy 123*(2), 444–496.

Borghans, L., J. J. Heckman, B. H. H. Golsteyn, and H. Meijers (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association 7*(2-3), 649–658.

Bramoulle, Y. and R. Kranton (2007). Risk-sharing networks. *Journal of Economic Behavior & Organization 64*(3-4), 275–294.

Brown, B. W. and W. K. Newey (2002). Generalized method of moments, efficient bootstrapping, and improved inference. *Journal of Business & Economic Statistics 20*(4), 507–517.

Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica 82*(4), 1273–1301.

Calvo-Armengol, A. (2004, March). Job contact networks. *Journal of Economic Theory 115*(1), 191–206.

Chandrasekhar, A. (2016). Econometrics of network formation. In *The Oxford Handbook of the Economics of Networks*. Oxford Handbook Online.

Cochran, W. G. (1952). The χ^2 test of goodness of fit. *The Annals of Mathematical Statistics 23*(3), 315–345.

Davidson, R. and J.-Y. Duclos (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica 68*(6), 1435–1464.

Dhrymes, P. J. (1978). *Mathematics for econometrics*. Springer.

Edwards, A. L. (1950). On 'The use and misuse of the chi-square test'- The case of the 2x2 contingency table. *Psychological bulletin 47*(4), 341.

Fafchamps, M. and S. Lund (2003). Risk-sharing networks in rural philippines. *Journal of development Economics 71*(2), 261–287.

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3), 253–264.

Goyal, S. (2012). *Connections: an introduction to the economics of networks*. Princeton University Press.

Graham, B. S. (2014). Methods of identification in social networks. Technical report, National Bureau of Economic Research.

Gumbel, E. J. (1943). On the reliability of the classical chi-square test. *The Annals of Mathematical Statistics 14*(3), 253–263.

Hahn, J. (1996). A note on bootstrapping generalized method of moments estimators. *Econometric Theory 12*(01), 187–197.

Hall, P. and J. L. Horowitz (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica 64*(4), 891–916.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society 46*(6), 1251–1271.

Hotelling, H. (1930). The consistency and ultimate distribution of optimum statistics. *Transactions of the American Mathematical Society 32*(4), 847–859.

Hsieh, C.-S. and L. F. Lee (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics 31*(2), 301–319.

Inoue, A. and M. Shintani (2006). Bootstrapping GMM estimators for time series. *Journal of Econometrics 133*(2), 531–555.

Jackson, M. O. (2014). Networks in the understanding of economic behaviors. *The Journal of Economic Perspectives 28*(4), 3–22.

Jackson, M. O. et al. (2008). *Social and economic networks.* Princeton University Press.

Jackson, M. O., T. Rodriguez-Barraquer, and X. Tan (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review 102*(5), 1857–97.

Leung, M. P. (2015). Two-step estimation of network-formation models with incomplete information. *Journal of Econometrics 188*(1), 182–195.

Lewis, D. and C. J. Burke (1949). The use and misuse of the chi-square test. *Psychological Bulletin 46*(6), 433.

Linton, O., E. Maasoumi, and Y.-J. Whang (2005). Consistent testing for stochastic dominance under general sampling schemes. *The Review of Economic Studies 72*(3), 735–765.

Liu, X. (2013). Estimation of a local-aggregate network model with sampled networks. *Economics Letters 118*(1), 243–246.

Macke, J. H., P. Berens, A. S. Ecker, A. S. Tolias, and M. Bethge (2009). Generating spike trains with specified correlation coefficients. *Neural Computation 21*(2), 397–423.

Mann, H. B. and A. Wald (1942). On the choice of the number of class intervals in the application of the chi square test. *The Annals of Mathematical Statistics 13*(3), 306–317.

McFadden, D. (1989). Testing for stochastic dominance. In *Studies in the Economics of Uncertainty*, pp. 113–134. Springer.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50*(302), 157–175.

Richmond, J. (1982). A general method for constructing simultaneous confidence intervals. *Journal of the American Statistical Association 77*(378), 455–460.

Rosenzweig, M. R. and O. Stark (1989). Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of Political Economy 97*(4), 905–926.

Rothschild, M. and J. E. Stiglitz (1970). Increasing risk: I. A definition. *Journal of Economic Theory 2*(3), 225 – 243.

Schorr, B. (1974). On the choice of the class intervals in the application of the chi-square test. *Statistics: A Journal of Theoretical and Applied Statistics 5*(4-5), 357–377.

Stoline, M. R. and H. K. Ury (1979). Tables of the studentized maximum modulus distribution and an application to multiple comparisons among means. *Technometrics 21*(1), pp. 87–93.

Tesfatsion, L. (1997). A trade network game with endogenous partner selection. In *Computational approaches to economic problems*, pp. 249–269. Springer.

Williams, C. A. (1950). The choice of the number and width of classes for the chi-square test of goodness of fit. *Journal of the American Statistical Association 45*(249), 77–86.

Yates, F. (1934). Contingency table involving small numbers and the chi-square test. *Journal of the Royal Statistical Society 1*(2), 217–235.