Consistent Testing for an Implication of Supermodular Dominance

Danbi Chung	Oliver Linton	Yoon-Jae Whang
Seoul National University	University of Cambridge	Seoul National University

Abstract

Supermodularity, or complementarity, is a popular concept in economics which can characterize many objective functions, including utility, social welfare, and production functions. Further, supermodular dominance captures a preference for greater interdependence among inputs of those functions, and it can be applied to examine which input set would produce higher expected utility, social welfare, or production. However, contrary to the profuse literature on supermodularity, to the best of our knowledge, there is no existing work on either testing or empirical analysis for supermodular dominance. In this paper, we propose a consistent test for a useful implication of supermodular dominance and suggest a correlation dominance testing for Gaussian random variables as a special case. The test is based on a novel bootstrap critical value, which has potentially enhanced power performance by exploiting the information on the contact set on which the null hypothesis is binding. We also conduct Monte Carlo simulations to explore the finite sample performance of our tests. We then apply our test to analyze two economic questions. We first investigate whether the interdependence of stock returns among major firms has increased after the COVID-19, and find evidence supporting this conjecture. We also compare the interdependence of patent citations depending on distance, where greater interdependence can imply greater expected social welfare effect. The results suggest that, in most cases, between-state citations seem to have greater interdependence than within-state citations, implying that lively interaction between firms across states might engender greater expected social welfare than knowledge spillover within a geographically confined area.

Reference Details

2134	Cambridge Working Papers in Economics
2021/15	Cambridge-INET Working Paper Series
Published	15 April 2021
Key Words	Supermodularity, Supermodular Dominance, Stochastic Dominance, Bootstrap, Contact Set, COVID-19, Patent Citation
JEL Codes	C63, D72, D83, D85, D91, L15
Websites	www.econ.cam.ac.uk/cwpe www.inet.econ.cam.ac.uk/working-papers

CONSISTENT TESTING FOR AN IMPLICATION OF SUPERMODULAR DOMINANCE

DANBI CHUNG, OLIVER LINTON, AND YOON-JAE WHANG

Abstract. Supermodularity, or complementarity, is a popular concept in economics which can characterize many objective functions, including utility, social welfare, and production functions. Further, supermodular dominance captures a preference for greater interdependence among inputs of those functions, and it can be applied to examine which input set would produce higher expected utility, social welfare, or production. However, contrary to the profuse literature on supermodularity, to the best of our knowledge, there is no existing work on either testing or empirical analysis for supermodular dominance. In this paper, we propose a consistent test for a useful implication of supermodular dominance and suggest a correlation dominance testing for Gaussian random variables as a special case. The test is based on a novel bootstrap critical value, which has potentially enhanced power performance by exploiting the information on the contact set on which the null hypothesis is binding. We also conduct Monte Carlo simulations to explore the finite sample performance of our tests. We then apply our test to analyze two economic questions. We first investigate whether the interdependence of stock returns among major firms has increased after the COVID-19, and find evidence supporting this conjecture. We also compare the interdependence of patent citations depending on distance, where greater interdependence can imply greater expected social welfare effect. The results suggest that, in most cases, between-state citations seem to have greater interdependence than within-state citations, implying that lively interaction between firms across states might engender greater expected social welfare than knowledge spillover within a geographically confined area.

Keywords: Supermodularity, Supermodular Dominance, Stochastic Dominance, Bootstrap, Contact Set, COVID-19, Patent Citation

Date: April 15, 2021.

1. Introduction

The concept of supermodularity has been widely used in economics to analyze how one agent's decision affects the incentive of the others. It is often called strategic complementarity, indicating that agents' strategies are complements to each other. In economics, many objective functions including utility, social welfare, and production functions, can be regarded as supermodular (or submodular[1]) functions. For instance, as Meyer and Strulovici (2017) pointed out, any financial loss function that is convex is supermodular in the individual losses. Another example includes a submodular production function, which appears when an increase in a firm's output has a negative effect on the outputs of the other firms. In general, supermodularity of a function is closely linked to positive interdependence among inputs.

Likewise, supermodular dominance, or supermodular ordering, is a general notion to capture preference for greater interdependence among inputs of those functions. It is an ordering between two sets of variables based on the ordering of the expected values of the variables evaluated with supermodular functions, details of which will be reviewed in Section 2. Thus, supermodular dominance can be applied to examine which input set would produce higher expected utility, social welfare, or production, a question of great interest in many areas of economics.

In this paper, we propose a formal test of an implication of supermodular dominance that is consistent against general alternatives. The test is based on a necessary condition of supermodular dominance, so that rejecting the test would imply that the underlying variables do not characterize supermodular ordering. We also suggest a test of correlation dominance, which is a special case of the supermodular dominance when the underlying variables are Gaussian. The limiting null distributions of our test statistics are shown to be non-pivotal and we propose a novel bootstrap critical value, which has a potentially enhanced power property by exploiting the information on the contact set on which the null hypothesis is binding.

We then investigate the finite sample performance of our tests by Monte Carlo simulations. The simulation designs include weakly dependent and strictly stationary data as well as

[1]A function f is submodular if $-f$ is supermodular.

independently and identically distributed multivariate normal data. In both cases, we confirm that our tests have reasonably good size and power performance under various settings.

Finally, we present two economic applications of our test. We first investigate whether the interdependence between the standardized stock returns of major firms have increased before and after the COVID-19 pandemic. Our tests suggest evidence in favor of this conjecture, on par with our intuition. We also compare the interdependence of patent citations based on distance, using patent citation data from the National Bureau of Economic Research (NBER), where greater interdependence can imply greater expected social welfare effect. The results suggest that, in most cases, between-state citations seem to have greater interdependence than within-state citations, indicating that lively interaction among firms across broad areas can engender higher expected social welfare than knowledge spillover, measured by patent citation, within a geographically confined area.

Literature Review. There is an extensive literature on supermodularity or complementarity in various branches of economics including game theory, macroeconomics, and comparative statics analysis. For instance, games with strategic complementarity, those wherein the best response of a player increases in another agent's strategy, gained interest starting from Topkis (1979), Vives (1985), and Vives (1990). Vives (1990) proposes a method to analyze Nash equilibria in noncooperative games utilizing lattice theory which best works in the presence of strategic complementarities. Vives (2005) also gives an introduction to the analysis of supermodular games and applies the result to the issues of industrial organization. Among the literature on comparative statics, Athey (2002) mentions that the robustness of comparative statics in economic theory is guaranteed by the (log-)supermodularity of primitive functions such as utility functions and probability distributions. Amir (2005) gives a non-technical survey.

In macroeconomics, Cooper and Haltiwanger (1996) provide theoretical propositions along with empirical evidence that, in the presence of complementarity, agents will show positive co-movement and have synchronized decisions, and aggregate shocks will be magnified and propagated. Acemoglu and Azar (2020) develop a model of endogenous production networks and show that supermodularity in all arguments of the production function implies the technology-price single-crossing condition, which implies that a positive technology shock or a reduction in distortions encourages technology adoption by all industries.

In finance, supermodularity plays an important role in providing general bounds on the price of stochastic annuities, Goovaerts and Dhaene (1999), and multiasset options such as basket or spread options, Rapuch and Roncalli (2004). In the Brownian motion (or Gaussian random variable) case it is known that the supermodular ordering of payoffs with identical marginals is equivalent to the ordering of the pairwise correlations, Muller and Scarsini (2000, Theorem 4.2). Kızıldemir and Privault (2015) and Kızıldemir (2017) showed that these conditions are also necessary and sufficient conditions for supermodular ordering of Poisson random vectors and extended the sufficiency results to a much larger family that captures the prototypical jump diffusion models widely used in the option pricing literature.

Relative to the profuse references on supermodularity, supermodular dominance has received little attention in the literature. Meyer and Strulovici (2017), in their pioneering research, study interdependence in economic settings using supermodular objective functions and suggest possible applications to various economic questions. However, there appears to be no existing statistical inference procedure available for supermodular dominance, partially because there is lack of the necessary and sufficient condition for the concept. Moreover, we have not been able to find any empirical studies that test for supermodularity or supermodular dominance using actual economic data. Our paper tries to fill this gap.

The paper is organized as follows. In Section 2, we review basic concepts and necessary conditions of supermodular dominance. Section 3 formulates the hypotheses of interest, introduces the test statistics, and defines bootstrap critical values. Section 4 discusses a test of correlation dominance for Gaussian random variables. Section 5 reports the Monte Carlo simulation results, and Section 6 presents two empirical applications. Section 7 is a conclusion. Finally, Appendix provides proofs of the main theorems.

2. Supermodular dominance

In this section, we review some basic concepts. A set $\mathcal{X} \subset \mathbb{R}^K$ is called a lattice if $x \wedge x' \in \mathcal{X}$ and $x \vee x' \in \mathcal{X}$ for any $x, x' \in \mathcal{X}$.[2]

[2] $x \wedge x'$ and $x \vee x'$ denote the element-wise minimum and element-wise maximum of two vectors x and x', respectively.

Definition 2.1. *A function $f : \mathcal{X} \to \mathbb{R}$ is supermodular on a lattice $\mathcal{X}$ if for any $x, x' \in \mathcal{X}$,*

$$f(x) + f(x') \leq f(x \wedge x') + f(x \vee x').$$

For a twice continuously differentiable f, supermodularity of f corresponds to non-negative cross-partial derivatives between all possible pairs of inputs, i.e.,

$$\frac{\partial^2 f(x_1, \ldots, x_K)}{\partial x_i \partial x_j} \geq 0 \; \forall \; i \neq j.$$

Let $X = (X_1, \ldots, X_K)^\intercal$ and $Y = (Y_1, \ldots, Y_K)^\intercal$ denote two random vectors in $\mathbb{R}^K$.

Definition 2.2. *Y supermodularly dominates X, denoted $X \leq_{sm} Y$, if*

$$Ef(X) \leq Ef(Y)$$

for all measurable, supermodular functions $f : \mathbb{R}^K \to \mathbb{R}$ for which the expectations exist.[3]

It is well known that, so far, there are no necessary and sufficient conditions found for supermodular dominance, but there are several implications at hand. The following lemma states some necessary conditions of the supermodular dominance proved by Bäuerle (1997):[4]

Lemma 2.1. *If $X \leq_{sm} Y$, then*

1) $X_i \overset{d}{=} Y_i$ *for each* $i = 1, \ldots, K$. (1)

2) $\max\{X_1, \ldots, X_K\} \geq_{st} \max\{Y_1, \ldots, Y_K\}$ *and* $\min\{X_1, \ldots, X_K\} \leq_{st} \min\{Y_1, \ldots, Y_K\}$.

3) $(X_{i_1}, \ldots, X_{i_k}) \leq_{sm} (Y_{i_1}, \ldots, Y_{i_k}) \; \forall \; 1 \leq i_1 < ... < i_k \leq K$.

Condition 1) implies that the variables in the supermodular dominance relationship have the same marginal distributions. Combining 2) and 3), we obtain

$$\begin{aligned} &\max\{X_{i_1}, \ldots, X_{i_k}\} \geq_{st} \max\{Y_{i_1}, \ldots, Y_{i_k}\} \text{ and} \\ &\min\{X_{i_1}, \ldots, X_{i_k}\} \leq_{st} \min\{Y_{i_1}, \ldots, Y_{i_k}\} \end{aligned} \tag{2}$$

[3]The usual first order stochastic dominance of Y over X, denoted $X \leq_{st} Y$, holds if the same inequality is satisfied for all measurable, *increasing* functions $f : \mathbb{R}^K \to \mathbb{R}$ for which the expectations exist.

[4]For any two random variables X and Y, we denote $X \overset{d}{=} Y$ when X and Y have the same marginal distribution.

for all $1 \leq i_1 < ... < i_k \leq K,\ k = 1, \ldots, K$, which is the main implication of our interest. That is, if there is supermodular dominance between two random vectors, the maximum of the elements of the supermodularly dominated random vector first order stochastically dominates that of the supermodularly dominating random vector for all possible combinations of the elements, and vice versa for the minimum.

Depending on the purpose of the research, one can test (2) for its own implication or test supermodular dominance by testing both (1) and (2) and see if both of them are satisfied. Since there are many existing tests for the equality of the distributions (1), e.g., the Kolmogorov-Smirnov two-sample test, below we shall focus on consistent testing for the property (2).

3. Testing for an implication of supermodular dominance

3.1. **Hypotheses of Interest and the Test Statistic.** In this section, we formulate the hypotheses of interest and define our test statistic.

Let $\mathcal{I} = \{(i_1, \ldots, i_k) \in \mathbb{Z}_+^k : 1 \leq i_1 < ... < i_k \leq K,\ k = 1, \ldots, K\}$ denote the set of k-tuples of positive integers. For each $\mathbf{i} \in \mathcal{I}$, let $F_{\mathbf{i}}^U$ and $G_{\mathbf{i}}^U$denote the distributions of $\max\{X_{i_1}, \ldots, X_{i_k}\}$ and $\max\{Y_{i_1}, \ldots, Y_{i_k}\}$, respectively. Similarly, we denote $F_{\mathbf{i}}^L$ and $G_{\mathbf{i}}^L$ to be the distributions of $\min\{X_{i_1}, \ldots, X_{i_k}\}$ and $\min\{Y_{i_1}, \ldots, Y_{i_k}\}$, respectively.

Let

$$D_{\mathbf{i}}^U(x) := F_{\mathbf{i}}^U(x) - G_{\mathbf{i}}^U(x); \quad D_{\mathbf{i}}^L(x) := G_{\mathbf{i}}^L(x) - F_{\mathbf{i}}^L(x). \tag{3}$$

The hypotheses of interest are as follows:

$$H_0 : D_{\mathbf{i}}^U(x) \leq 0 \text{ and } D_{\mathbf{i}}^L(x) \leq 0 \text{ for all } x \in \mathcal{X} \text{ and } \mathbf{i} \in \mathcal{I}, \tag{4}$$

$$H_1 : \text{Negation of } H_0.$$

In this formulation, rejecting the null hypothesis would imply that Y does not supermodularly dominate X, while failing to reject the null implies that there is not enough evidence against the necessary condition for the supermodular dominance of Y over X.

Consider the population quantity:

$$d_0 := \sum_{\mathbf{i}\in\mathcal{I}} \int_{\mathcal{X}} \{[D_{\mathbf{i}}^U(x)]_+^2 + [D_{\mathbf{i}}^L(x)]_+^2\} dx. \tag{5}$$

Then, the hypotheses of interest can be equivalently stated as

$$H_0 : d_0 = 0 \text{ vs. } H_0 : d_0 > 0. \tag{6}$$

Let $\mathbf{X}_t = (X_{1,t}, \ldots, X_{K,t})^\intercal$ and $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{K,t})^\intercal$ for $t = 1, \ldots, T$, where $\{X_{i,t} : t = 1, \ldots, T\}$ and $\{Y_{i,t} : t = 1, \ldots, T\}$ denote the sample realizations of X_i and Y_i, respectively, for $i = 1, \ldots, K$. We impose the following assumptions:

Assumption 3.1. *(a) $\{(\mathbf{X}_t, \mathbf{Y}_t) : t \geq 1\}$ is a strictly stationary and strong mixing sequence with coefficients $\{\alpha(s) : s \geq 1\}$ satisfying $\sum_{s=1}^{\infty} s^{Q-1}\alpha(s)^{\gamma/(Q+\gamma)} < \infty$ for some even integer $Q > 2 + \gamma$ and some $\gamma > 0$. (b) The distributions of $(X_{i_1,t}, \ldots, X_{i_k,t})$ and $(Y_{i_1,t}, \ldots, Y_{i_k,t})$ have bounded densities with respect to Lebesgue measure for each $(i_1, \ldots, i_k) \in \mathcal{I}$.*

Our test statistic is based on the sample analogue of (5):

$$S_T := T \sum_{\mathbf{i}\in\mathcal{I}} \int_{\mathcal{X}} \{[\hat{D}_{\mathbf{i}}^U(x)]_+^2 + [\hat{D}_{\mathbf{i}}^L(x)]_+^2\} dx,$$

where $\hat{D}_{\mathbf{i}}^U(x) := \hat{F}_{\mathbf{i}}^U(x) - \hat{G}_{\mathbf{i}}^U(x)$ and $\hat{D}_{\mathbf{i}}^L(x) := \hat{G}_{\mathbf{i}}^L(x) - \hat{F}_{\mathbf{i}}^L(x)$, with $\hat{F}_{\mathbf{i}}^U$, $\hat{G}_{\mathbf{i}}^U$, $\hat{F}_{\mathbf{i}}^L$, and $\hat{G}_{\mathbf{i}}^L$ denoting the empirical distribution functions of $\max\{X_{i_1,t}, \ldots, X_{i_k,t}\}$, $\max\{Y_{i_1,t}, \ldots, Y_{i_k,t}\}$, $\min\{X_{i_1,t}, \ldots, X_{i_k,t}\}$, and $\min\{Y_{i_1,t}, \ldots, Y_{i_k,t}\}$, respectively.

The limit distribution of the test statistic under the null hypothesis can be characterized using fact that

$$\sqrt{T}\left(\hat{D}_{\mathbf{i}}(\cdot) - D_{\mathbf{i}}(\cdot)\right) \Rightarrow \nu_{\mathbf{i}}(\cdot)$$

jointly for all $\mathbf{i} \in \mathcal{I}$, where $\hat{D}_{\mathbf{i}}(\cdot) = (\hat{D}_{\mathbf{i}}^U(\cdot), \hat{D}_{\mathbf{i}}^L(\cdot))^\intercal$, $D_{\mathbf{i}}(\cdot) = (D_{\mathbf{i}}^U(\cdot), D_{\mathbf{i}}^L(\cdot))^\intercal$, and $\nu_{\mathbf{i}}(\cdot) = (\nu_{\mathbf{i}}^U(\cdot), \nu_{\mathbf{i}}^L(\cdot))^\intercal$ is a mean zero Gaussian process with covariance kernel $C(x_1, x_2) = E\nu_{\mathbf{i}}(x_1)\nu_{\mathbf{i}}(x_2)^\intercal$ (Lemma A1 in Appendix).

A decision rule of the test would be rejecting H_0 if $S_T > c_{T,\alpha}$ where $c_{T,\alpha}$ is a critical value that will be discussed in the next section. Since the asymptotic null distribution of the test statistic depends on the true DGPs, we suggest resampling methods to compute the critical values.

3.2. **Bootstrap Critical Values.** We assume that the observations are weakly dependent and strictly stationary with arbitrary contemporaneous dependence. To compute the critical values, we therefore consider the stationary bootstrap procedure (Politis and Romano, 1994a) applicable to time series settings.

Let $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_T)^\intercal$ denote the $T \times 2K$ matrix of the original sample, where $\mathbf{W}_t = (\mathbf{X}_t^\intercal, \mathbf{Y}_t^\intercal)^\intercal \in \mathbb{R}^{2K}$, $\mathbf{X}_t = (X_{1,t}, \ldots, X_{K,t})^\intercal$ and $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{K,t})^\intercal$ for $t = 1, \ldots, T$. Let $B_{i,l} := (\mathbf{W}_i, \mathbf{W}_{i+1}, \ldots, \mathbf{W}_{i+l-1})^\intercal$ denote the block (i.e., submatrix) of $\mathbf{W}$ that starts from the i-th row and has a total of l rows. When $j > T$, we let $\mathbf{W}_j = \mathbf{W}_{j(mod\,T)}$ and $\mathbf{W}_0 = \mathbf{W}_T$.

Let $\{L_i : i \geq 1\}$ denote a sequence of block lengths,[5] which are iid random variables independent of $\mathbf{W}$ having a geometric distribution with a scalar parameter $p \in (0, 1)$, i.e.,

$$P(L_i = l) = (1 - p)^{l-1} p \text{ for } l = 1, 2, \ldots; i \geq 1.$$

We assume that the parameter p satisfies the following growth condition:

Assumption 3.2. $p + (\sqrt{T}p)^{-1} \to \infty$.

Let $\{I_i : i \geq 1\}$ be iid random variables independent of $\mathbf{W}$ and L_i's, having a uniform distribution on $\{1, \ldots, T\}$, i.e.,

$$P(I_i = s) = \frac{1}{T} \text{ for } s = 1, 2, \ldots, T; i \geq 1.$$

Then we take $\mathbf{W}^* = \left(B_{I_1,L_1}^\intercal, B_{I_2,L_2}^\intercal, \ldots, B_{I_k,L_k}^\intercal\right)^\intercal := (\mathbf{W}_1^*, \ldots, \mathbf{W}_T^*)^\intercal$ as our bootstrap sample, where $\mathbf{W}_t^* := (\mathbf{X}_t^{*\intercal}, \mathbf{Y}_t^{*\intercal})^\intercal$, $\mathbf{X}_t^* = (X_{1,t}^*, \ldots, X_{K,t}^*)^\intercal$ and $\mathbf{Y}_t^* = (Y_{1,t}^*, \ldots, Y_{K,t}^*)^\intercal$, i.e., the first L_1 rows of $\mathbf{W}^*$ are B_{I_1,L_1}, the next L_2 rows are B_{I_2,L_2}, and so on. This procedure is stopped when total T observations are generated for each column, and we discard the remaining. Politis and Romano (1994a) show that $\mathbf{W}^*$ is stationary, conditional on $\mathbf{W}$.

The bootstrap version of the empirical distribution $\hat{F}_{\mathbf{i}}^U$ is given by

$$\hat{F}_{\mathbf{i}}^{U*}(x) = \frac{1}{T} \sum_{t=1}^{T} 1\left(\max\{X_{i_1,t}^*, \ldots, X_{i_k,t}^*\} \leq x\right).$$

Likewise, $\hat{G}_{\mathbf{i}}^{U*}$, $\hat{F}_{\mathbf{i}}^{L*}$ and $\hat{G}_{\mathbf{i}}^{L*}$ can be defined.

[5] If the data are independently and identically distributed, then we may take $L_i = 1$, which essentially leads to the standard nonparametric bootstrap.

One may consider the following (re-centered) bootstrap test statistic:

$$\hat{S}^*_{T,LF} = T \sum_{\mathbf{i} \in \mathcal{I}} \int_{\mathcal{X}} \{[\hat{D}^{U*}_{\mathbf{i}}(x) - \hat{D}^{U}_{\mathbf{i}}(x)]^2_+ + [\hat{D}^{L*}_{\mathbf{i}}(x) - \hat{D}^{L}_{\mathbf{i}}(x)]^2_+\} dx, \tag{7}$$

where $\hat{D}^{U*}_{\mathbf{i}}(x) := \hat{F}^{U*}_{\mathbf{i}}(x) - \hat{G}^{U*}_{\mathbf{i}}(x)$ and $\hat{D}^{L*}_{\mathbf{i}}(x) := \hat{G}^{L*}_{\mathbf{i}}(x) - \hat{F}^{L*}_{\mathbf{i}}(x)$. Then, the *LFC-based bootstrap critical value* $c^*_{T,\alpha,LF}$ is defined to be the $(1-\alpha)$ quantile of the bootstrap distribution of $\hat{S}^{*,LF}_T$.

However, the test statistics based on the LFC-based bootstrap critical value can be too conservative, mainly because the null hypothesis (4) may comprise of a large number of inequalities restrictions (especially when K is large) and only a few of them may be binding. In this paper, we propose a bootstrap critical value with potentially enhanced power properties by exploiting the information on the *contact set* over which the inequalities are binding. [6].

For positive sequences $\{(a_{T,1}, a_{T,2}) : T \geq 1\}$, let

$$\mathcal{B}_{\mathbf{i},1} = \{x \in \mathcal{X} : |D^U_{\mathbf{i}}(x)| \leq a_{T,1} \text{ and } D^L_{\mathbf{i}}(x) < -a_{T,2}\}, \tag{8}$$

$$\mathcal{B}_{\mathbf{i},2} = \{x \in \mathcal{X} : |D^L_{\mathbf{i}}(x)| \leq a_{T,1} \text{ and } D^U_{\mathbf{i}}(x) < -a_{T,2}\}, \tag{9}$$

$$\mathcal{B}_{\mathbf{i},3} = \{x \in \mathcal{X} : |D^U_{\mathbf{i}}(x)| \leq a_{T,1} \text{ and } |D^L_{\mathbf{i}}(x)| \leq a_{T,1}\}, \tag{10}$$

denote the (population) contact sets, where $D^U_{\mathbf{i}}$ and $D^L_{\mathbf{i}}$ are as defined in (3).

Lemma 3.1. *Suppose that Assumption 3.1 holds. Suppose further that $a_{T,1}, a_{T,2}$ are positive sequences satisfying Assumption 3.3 below. Then, under the null hypothesis H_0, we have*

$$P\left\{S_T = T \sum_{\mathbf{i} \in \mathcal{I}} \left(\int_{\mathcal{B}_{\mathbf{i},1}} [\hat{D}^U_{\mathbf{i}}(x)]^2_+ dx + \int_{\mathcal{B}_{\mathbf{i},2}} [\hat{D}^L_{\mathbf{i}}(x)]^2_+ dx + \int_{\mathcal{B}_{\mathbf{i},3}} \{[\hat{D}^U_{\mathbf{i}}(x)]^2_+ + [\hat{D}^L_{\mathbf{i}}(x)]^2_+\} dx \right) \right\} \to 1.$$

Lemma 3.1 shows that, under the null hypothesis, the test statistic S_T can be approximated by the sum of the integrals with domains restricted to the contact sets with probability that goes to one as the sample size goes to infinity. This suggests that we may consider a bootstrap procedure that mimics the approximate representation of S_T in Lemma 3.1.

[6] See Linton, Song and Whang (2010) for the original definition of the contact set in the context of testing for stochastic dominance.

Let

$$\hat{\mathcal{B}}_{\mathbf{i},1} = \{x \in \mathcal{X} : |\hat{D}_{\mathbf{i}}^{U}(x)| < \hat{a}_T \text{ and } \hat{D}_{\mathbf{i}}^{L}(x) < -\hat{a}_T\},$$
$$\hat{\mathcal{B}}_{\mathbf{i},2} = \{x \in \mathcal{X} : |\hat{D}_{\mathbf{i}}^{L}(x)| < \hat{a}_T \text{ and } \hat{D}_{\mathbf{i}}^{U}(x) < -\hat{a}_T\},$$
$$\hat{\mathcal{B}}_{\mathbf{i},3} = \{x \in \mathcal{X} : |\hat{D}_{\mathbf{i}}^{U}(x)| < \hat{a}_T \text{ and } |\hat{D}_{\mathbf{i}}^{L}(x)| < \hat{a}_T\}.$$

denote the *estimated* contact sets, where $\{\hat{a}_T : T \geq 1\}$ is a positive (possibly stochastic) sequence that satisfies the following assumption:

Assumption 3.3. *For each $T \geq 1$, there exist non-stochastic sequences $a_{T,1}, a_{T,2} > 0$ such that $a_{T,1} \leq a_{T,2}$ and*

$$P\left\{a_{T,1} \leq \hat{a}_T \leq a_{T,2}\right\} \to 1 \textit{ and } a_{T,1}\sqrt{T} + a_{T,2}^{-1} \to \infty \textit{ as } T \to \infty.$$

Then, our modified bootstrap statistic is defined by

$$S_T^* = T\sum_{\mathbf{i}\in\mathcal{I}}\left(\int_{\hat{\mathcal{B}}_{\mathbf{i},1}}[\hat{D}_{\mathbf{i}}^{U*}(x) - \hat{D}_{\mathbf{i}}^{U}(x)]_+^2 dx + \int_{\hat{\mathcal{B}}_{\mathbf{i},2}}[\hat{D}_{\mathbf{i}}^{L*}(x) - \hat{D}_{\mathbf{i}}^{L}(x)]_+^2 dx \right.$$
$$\left. + \int_{\hat{\mathcal{B}}_{\mathbf{i},3}}\{[\hat{D}_{\mathbf{i}}^{U*}(x) - \hat{D}_{\mathbf{i}}^{U}(x)]_+^2 + [\hat{D}_{\mathbf{i}}^{L*}(x) - \hat{D}_{\mathbf{i}}^{L}(x)]_+^2\}dx\right). \tag{11}$$

It can be easily seen that the modified bootstrap statistic 11 is generally smaller than the LFC based bootstrap statistic 7, because the estimated contact sets are subsets of $\mathcal{X}$. Therefore, it allows employing smaller critical values that the LFC based critical values without sacrificing the size, which in turn enhances the power performance of the test.

Let S_T^* be the $(1-\alpha)$ quantile from the bootstrap distribution S_T^*. In some interior cases of the null hypothesis, both the original test statistic S_T and the modified bootstrap statistic S_T^* may degenerate to zero as the sample size increases. To control the size, therefore, we suggest taking the bootstrap critical value as the maximum of S_T^* and an arbitrary small positive constant:

$$c_{T,\alpha,\eta}^* = \max\{c_{T,\alpha}^*, \eta\},$$

where $\eta = 10^{-6}$ is a small fixed number. The decision rule of the test is to reject the null hypothesis if the test statistic S_T exceeds the bootstrap critical value $c_{T,\alpha,\eta}^*$.

The following theorem shows that our test has asymptotically correct size under the null hypothesis[7] and is consistent against the fixed alternative hypothesis.

Theorem 3.2. *Suppose that Assumptions 3.1, 3.2, and 3.3 hold. (i) Then, under the null hypothesis H_0,*

$$\limsup_{T\to\infty} P\left\{S_T > c^*_{T,\alpha,\eta}\right\} \le \alpha.$$

(ii) Under the alternative hypothesis H_1, we have

$$P\left\{S_T > c^*_{T,\alpha,\eta}\right\} \to 1.$$

4. Testing for correlation dominance

In this section, we consider a test of correlation dominance, which is a special case of the supermodular dominance when the random variables are Gaussian.

Let $X = (X_1, \ldots, X_K)^\intercal \in \mathbb{R}^K$ and $Y = (Y_1, \ldots, Y_K)^\intercal \in \mathbb{R}^K$ be two Gaussian random vectors with the same marginal distributions $X_i \overset{d}{=} Y_i,\ i = 1, \ldots, K$. Then, by Müller and Scarsini (2000, Theorem 4.2), we have $X \le_{sm} Y$ if and only if,

$$\text{cov}(X_i, X_j) \le \text{cov}(Y_i, Y_j)\ \ \forall i, j = 1, .., K;\ i \ne j. \tag{12}$$

We say that Y *correlation dominates* X, or $X \le_{cd} Y$, when (12) holds.

Let $\{X_{i,t} : t = 1, \ldots, T\}$ and $\{Y_{i,t} : t = 1, \ldots, T\}$ be the sample realizations of X_i and Y_i, respectively, for $i = 1, \ldots, K$. To test the correlation dominance (12), we consider the following test statistic:

$$C_T = \sqrt{T} \max_{1\le i<j\le K} \{\hat{\Sigma}_{i,j}(X) - \hat{\Sigma}_{i,j}(Y)\},$$

where $\hat{\Sigma}_{i,j}(X) = \frac{1}{T}\sum_{t=1}^T (X_{i,t} - \bar{X}_i)(X_{j,t} - \bar{X}_j), \bar{X}_i = \frac{1}{T}\sum_{t=1}^T X_{i,t}$, and $\hat{\Sigma}_{i,j}(Y) = \frac{1}{T}\sum_{t=1}^T (Y_{i,t} - \bar{Y}_i)(Y_{j,t} - \bar{Y}_j), \bar{Y}_i = \frac{1}{T}\sum_{t=1}^T Y_{i,t}$ for $i, j = 1, \ldots, K$. Using a standard argument, it can be seen that the asymptotic distribution of the test statistic C_T under (the least favorable case of) the null hypothesis (12) is a maximum of Gaussian distributions.

[7] Under additional regularity conditions, we may show that the test has asymptotically correct size *uniformly* under the null hypothesis. For brevity, we do not discuss the result in this paper.

We consider the following bootstrap test statistic based on the bootstrap sample $\mathbf{W}^*$ defined in Section 3.2:

$$C_T^* = \sqrt{T} \max_{1 \le i < j \le K} \{\hat{\Sigma}_{i,j}^*(X) - \hat{\Sigma}_{i,j}^*(Y) - (\hat{\Sigma}_{i,j}(X) - \hat{\Sigma}_{i,j}(Y))\}, \tag{13}$$

where $\hat{\Sigma}_{i,j}^*(X) = \frac{1}{T}\sum_{t=1}^T (X_{i,t}^* - \bar{X}_i^*)(X_{j,t}^* - \bar{X}_j^*)$, $\bar{X}_i^* = \frac{1}{T}\sum_{t=1}^T X_{i,t}^*$ and $\hat{\Sigma}_{i,j}^*(Y)$ is defined similarly. The bootstrap critical value with nominal significance level α is defined to be the $(1-\alpha)$ quantile of the bootstrap distribution of C_T^*.[8]

5. Monte Carlo simulations

In this section, we conduct Monte Carlo simulations to explore the finite sample performance of our test. In our simulations, we consider both iid data and serially dependent data.

5.1. **IID data.** We randomly generate T samples of vectors of dimension $K = 3$ from the following multivariate normal (MVN) distributions:

$$\begin{aligned} X &\sim MVN(\mathbf{0}, \Sigma_X), \\ Y &\sim MVN(\mathbf{0}, \Sigma_Y). \end{aligned} \tag{14}$$

We consider several cases to examine the size and power of our test. In Case 1, both X and Y follow the multivariate standard normal distribution (14) with $\Sigma_X = \Sigma_Y = I_K$. In other cases, we vary the covariances to modify the dependence structure. In Case 2, the off-diagonals of Σ_X are set to a positive constant in $\{0.1, 0.3, 0.5\}$, so that X supermodularly dominates Y. In Case 3, the off-diagonals of Σ_X move between ± 0.2 or ± 0.4 in an orderly fashion. In both Cases 2 and 3, we expect the rejection rates to increase as the covariances differ greater from 0. Case 4 varies from Case 2 in that some of the off-diagonals of Σ_X are set to zero. This mimics the situation where there is a non-negligible contact set, i.e., the tangent area between the null and alternative hypothesis. Therefore, we expect to reject the null hypothesis in Case 2 to 4. Finally, in Case 5, the covariances of Y are equal to 0.35 so

[8] It would also be possible to define a modified bootstrap test statistic utilizing the information on the binding inequalities, similar to (11). Since the correlation dominance test can be used as a preliminary screening tool and such an extension complicates the exposition, we do not discuss the case in this paper.

that Y has strict supermodular dominance over X, which characterizes an interior case of the null hypothesis. In sum, we consider

$$\Sigma_X^{(2)} = \begin{pmatrix} 1 & x & x \\ x & 1 & x \\ x & x & 1 \end{pmatrix}, \Sigma_X^{(3)} = \begin{pmatrix} 1 & y & -y \\ y & 1 & y \\ -y & y & 1 \end{pmatrix}, \Sigma_X^{(4)} = \begin{pmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \Sigma_Y^{(5)} = \begin{pmatrix} 1 & 0.35 & 0.35 \\ 0.35 & 1 & 0.35 \\ 0.35 & 0.35 & 1 \end{pmatrix}$$

with $x = 0.1, 0.3, 0.5$, and $y = 0.2, 0.4$, where the superscripts denote the corresponding cases. Σ_X and Σ_Y in other cases are identity matrices.

5.1.1. *Simulation Results.* We take $B = 500$ bootstrap repetitions and $R = 1,000$ simulation replications, and consider the sample size $T \in \{100, 500\}$. Since we are considering a multivariate standard normal distribution along with its variations, a grid of values is selected on the interval [-2.5, 2.5] such that $-2.5 = t_0 < t_1 < ... < t_n = 2.5$. The number of gridpoints is fixed at $n = 40$.

We take the nominal significance level $\alpha = 0.05$. As the tuning parameter for the contact set approach, we take $a_T = a_0 T^{-1/2} \log(\log(T))$. In the simulations, we consider $a_0 \in \{0.6, \ldots, 1.0\}$ to see the robustness of the results with respect to the choice of the tuning parameter constant.

Table 1 summarizes the simulation results. BS and BS_C refer to the tests with the LFC-based bootstrap critical value $c^*_{T,\alpha,LF}$ and the contact set-based bootstrap critical value $c^*_{T,\alpha,\eta}$, respectively. In Case 1, the rejection probabilities are close to the nominal significance level, confirming that our test has reasonably good size in finite samples. In Case 2, as expected, the rejection rates increase as $x = (\Sigma_X)_{ij}$ changes from 0.1 to 0.5, or as the sample size T increases. In Case 3, we have a similar finding: the power increases as the distance between the null and alternative models increases or the sample size increases. In Case 4, we find that the contact set based bootstrap critical values give significantly higher power than the LFC-based critical values, which is consistent with our theoretical results. In Case 5, as expected, the rejection probabilities of BS_C are closer to the nominal size than BS, which implies that BS_C test is asymptotically similar over a wider set of null distributions than BS.

Table 1. Rejection Probabilities of S_T: IID Data

T	$a_0 = 0.6$				$a_0 = 0.7$			
	100		500		100		500	
	BS	BS_C	BS	BS_C	BS	BS_C	BS	BS_C
Case 1	0.054	0.124	0.041	0.079	0.048	0.091	0.044	0.077
Case 2: $x = 0.1$	0.094	0.197	0.183	0.306	0.099	0.186	0.205	0.282
Case 2: $x = 0.3$	0.330	0.544	0.988	0.999	0.310	0.463	0.991	0.997
Case 2: $x = 0.5$	0.829	0.950	1.000	1.000	0.814	0.914	1.000	1.000
Case 3: $y = 0.2$	0.075	0.177	0.246	0.374	0.090	0.162	0.244	0.331
Case 3: $y = 0.4$	0.218	0.431	0.964	0.995	0.211	0.358	0.961	0.993
Case 4	0.205	0.378	0.917	0.969	0.217	0.355	0.916	0.961
Case 5	0.010	0.037	0.002	0.011	0.011	0.042	0.001	0.013
T	$a_0 = 0.8$				$a_0 = 0.9$			
	100		500		100		500	
	BS	BS_C	BS	BS_C	BS	BS_C	BS	BS_C
Case 1	0.049	0.088	0.046	0.070	0.056	0.084	0.047	0.061
Case 2: $x = 0.1$	0.078	0.123	0.170	0.211	0.088	0.120	0.192	0.206
Case 2: $x = 0.3$	0.319	0.405	0.987	0.991	0.319	0.383	0.989	0.991
Case 2: $x = 0.5$	0.832	0.899	1.000	1.000	0.806	0.856	1.000	1.000
Case 3: $y = 0.2$	0.086	0.125	0.212	0.278	0.105	0.142	0.221	0.253
Case 3: $y = 0.4$	0.185	0.282	0.968	0.986	0.223	0.292	0.965	0.978
Case 4	0.190	0.293	0.890	0.927	0.189	0.260	0.917	0.948
Case 5	0.009	0.020	0.000	0.011	0.010	0.021	0.003	0.007
T	$a_0 = 1.0$							
	100		500					
	BS	BS_C	BS	BS_C				
Case 1	0.047	0.059	0.054	0.061				
Case 2: $x = 0.1$	0.089	0.112	0.196	0.200				
Case 2: $x = 0.3$	0.330	0.361	0.986	0.987				
Case 2: $x = 0.5$	0.831	0.862	1.000	1.000				
Case 3: $y = 0.2$	0.079	0.101	0.205	0.223				
Case 3: $y = 0.4$	0.229	0.263	0.973	0.976				
Case 4	0.173	0.204	0.931	0.943				
Case 5	0.007	0.018	0.004	0.009				

Table 2. Rejection Probabilities of C_T : IID Data

T	100	250	500
Case 1	0.076	0.070	0.052
Case 2: $x = 0.1$	0.206	0.403	0.633
Case 2: $x = 0.3$	0.871	0.999	1.000
Case 2: $x = 0.5$	0.999	1.000	1.000
Case 3: $y = 0.2$	0.463	0.846	0.982
Case 3: $y = 0.4$	0.980	1.000	1.000
Case 4	0.955	1.000	1.000
Case 5	0	0	0

We next provide the simulation results of the correlation dominance test C_T. Table 2 summarizes the rejection probabilities. Overall, the test has reasonably good size and power performances in finite samples.

5.2. **Dependent data.** We next consider the cases where observations for X and Y are mutually and serially dependent. For this purpose, we generate $\mathbf{W}_t = (X_{1,t}, \ldots, X_{K,t}, Y_{1,t}, \ldots, Y_{Kt})^\intercal \in \mathbb{R}^{2K}$ as follows:

$$\mathbf{W}_t = A\mathbf{W}_{t-1} + \epsilon_t, \quad t = 1, \ldots T,$$

with $\mathbf{W}_0 = \mathbf{0}$, where $A = \text{diag}\{\rho, \ldots, \rho\}$ is a $2K \times 2K$ diagonal matrix, and ϵ_i is a $2K \times 1$ vector randomly generated from the multivariate normal distribution with mean $\mathbf{0}$ and variance-covariance matrix $\Sigma_\epsilon = (\gamma_{ij})$ with $\gamma_{ii} = 1$.

To investigate the size performance of the test, we set $\rho = 0.3$ and take $\gamma_{ij} = 0.2$ for all $i \neq j$. On the other hand, for the power performance, we vary the values of $\gamma_{cov} := \gamma_{ij},\ i \neq j$, so that the covariances between the elements of $\mathbf{X}_t = (X_{1,t}, ..., X_{K,t})^\intercal$ are greater than those of $\mathbf{Y}_t = (Y_{1,t}, ..., Y_{K,t})^\intercal$. The extent to which we deviate from the null hypothesis of supermodular (and correlation) dominance is characterized by the parameter $\gamma_{cov} \in \{0.3, 0.5\}$.[9] For other parameters, we take B (number of bootstrap repetition) = 500, R (number of simulation repetition) = 1000, $T \in \{100, 250, 500\}$, $a_0 \in \{1.0, 1.1, 1.2\}$ and $1/p \in \{5, 15\}$.

Table 3 summarizes the size and the power of the test S_T. Overall, it shows that our test has reasonably good performance under various settings and the test BS_C is more powerful than BS especially when the value of a_0 is smaller, as expected.

Tabel 4 summarizes the size and the power of the test C_T for the same parameter setting. As before, the finite sample performance of the correlation test seems reasonably good.

6. Applications

In this section, we apply our tests to two empirical questions. Since we will consider time series data, we use the stationary bootstrap to obtain the bootstrap critical values.

6.1. **Stock returns of major firms before and after the COVID-19.** Big Tech, also known as the Tech Giants or the Big Five are the largest and most dominant companies in the information technology industry of the United States: Amazon (AMZN), Apple (AAPL), Alphabet (GOOG), Facebook (FB), and Microsoft (MSFT). For the last ten years, these five have been the most valuable public companies globally, with market capitalizations reaching

[9] In specific, we partition the variance covariance matrix into four rectangles and set the off-diagonals of the upper left partition as γ_{cov}, that of the lower right partition as 0.1, and the rest as 0.05.

Table 3. Rejection Probabilities of the test S_T: Dependent Data

$1/p = 5$									
a_0=1.0	H_0			H_1: $\gamma_{cov} = 0.3$			H_1: $\gamma_{cov} = 0.5$		
T	100	250	500	100	250	500	100	250	500
BS	0.086	0.066	0.07	0.213	0.302	0.505	0.517	0.912	1
BS_C	0.12	0.08	0.078	0.253	0.327	0.539	0.555	0.929	1
a_0=1.1	H_0			H_1: $\gamma_{cov} = 0.3$			H_1: $\gamma_{cov} = 0.5$		
T	100	250	500	100	250	500	100	250	500
BS	0.086	0.063	0.072	0.211	0.31	0.562	0.531	0.905	1
BS_C	0.104	0.068	0.078	0.245	0.331	0.575	0.572	0.912	1
a_0=1.2	H_0			H_1: $\gamma_{cov} = 0.3$			H_1: $\gamma_{cov} = 0.5$		
T	100	250	500	100	250	500	100	250	500
BS	0.08	0.071	0.073	0.204	0.309	0.527	0.527	0.919	1
BS_C	0.086	0.077	0.074	0.223	0.313	0.529	0.55	0.921	1
$1/p = 15$									
a_0=1.0	H_0			H_1: $\gamma_{cov} = 0.3$			H_1: $\gamma_{cov} = 0.5$		
T	100	250	500	100	250	500	100	250	500
BS	0.113	0.073	0.062	0.264	0.293	0.51	0.565	0.9	1
BS_C	0.134	0.087	0.07	0.293	0.319	0.549	0.617	0.921	1
a_0=1.1	H_0			H_1: $\gamma_{cov} = 0.3$			H_1: $\gamma_{cov} = 0.5$		
T	100	250	500	100	250	500	100	250	500
BS	0.123	0.087	0.059	0.247	0.282	0.531	0.614	0.884	0.999
BS_C	0.14	0.099	0.064	0.276	0.304	0.551	0.652	0.896	1
a_0=1.2	H_0			H_1: $\gamma_{cov} = 0.3$			H_1: $\gamma_{cov} = 0.5$		
T	100	250	500	100	250	500	100	250	500
BS	0.124	0.061	0.052	0.274	0.299	0.52	0.59	0.893	0.998
BS_C	0.132	0.063	0.055	0.28	0.303	0.525	0.606	0.9	0.998

Table 4. Rejection Probabilities of the test C_T: Dependent Data

$1/p$=5			
T	H_0	H_1: $\gamma_{cov} = 0.3$	H_1: $\gamma_{cov} = 0.5$
100	0.076	0.670	0.994
250	0.058	0.929	1.000
500	0.057	0.998	1.000
$1/p$=15			
T	H_0	H_1: $\gamma_{cov} = 0.3$	H_1: $\gamma_{cov} = 0.5$
100	0.113	0.737	0.995
250	0.081	0.941	1.000
500	0.069	0.998	1.000

over \$2.0 trillion. They now collectively represent 25% of the market value of the S&P500 index.

The COVID-19 pandemic has changed the world economy. Many stock markets around the world suffered big drops in value at various points during the pandemic, but Big Tech seems to keep on growing in importance, since many more transactions are taking place online rather than face-to-face. We investigate how the Big Five stock prices fared during

2019 and 2020. According to the WHO, the pandemic started on December 31st 2019, and so we interpret 2020 as pandemic related and 2019 as pre-pandemic. Although the US did not suffer many cases until March 2020, the Big Five receive much of their earnings from international sources and so the calendar year division seems reasonable.

We consider their daily closing prices during 2019 and 2020. We calculate the return matrices $R_{2019}$and $R_{2020}$with 254×5 and 178×5 dimensions. Below we report the sample moments of the daily returns from the two years.

There seems to be some indication that the mean returns have increased in 2020, although the volatility has also increased. There does not seem to be a common pattern in the behavior of the higher cumulants. We test whether the means have changed significantly using the Hotelling statistic

$$H = (\bar{R}_{2020} - \bar{R}_{2019})'(\frac{C_{2020}}{T_2} + \frac{C_{2019}}{T_1})^{-1}(\bar{R}_{2020} - \bar{R}_{2019}),$$

where $T_i,\ i = 1, 2$is the number of observations in each year. The value is 2.773, which is not significant compared with the critical value from $\mathcal{X}^2(5)$. We also considered the standard event study methodology, where we compute abnormal return for 2020 as $AR_{k,t} = R_{k,t} - \hat{\mu}_{2019,t}$ and the cumulative abnormal return as $\sum_{s=1}^{t} AR_{k,s}$. To set the confidence bars we used the 2020 standard deviations. Figure 1 shows that there was an early negative effect of the pandemic on the stock returns of each company but they mostly stayed within the confidence bands.

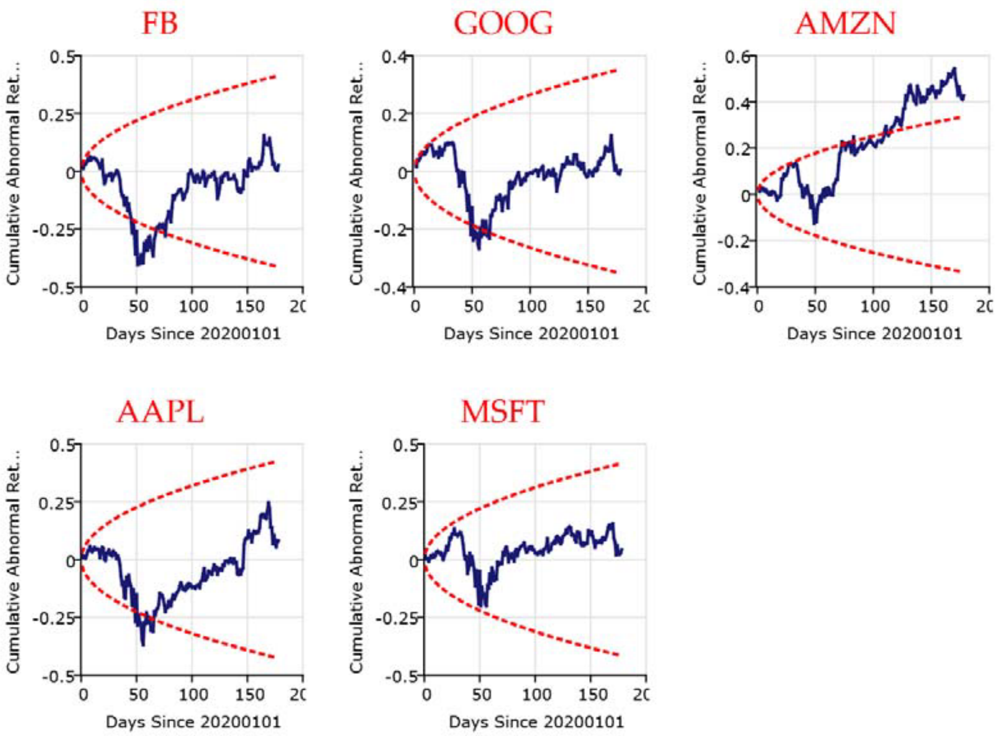
The sample correlation matrices for $X = R_{2019}$[10] and $Y = R_{2020}$are, respectively,

$$\hat{\sigma}_X = \begin{pmatrix} 1.000 & 0.578 & 0.493 & 0.509 & 0.548 \\ 0.578 & 1.000 & 0.685 & 0.545 & 0.741 \\ 0.493 & 0.685 & 1.000 & 0.549 & 0.675 \\ 0.509 & 0.545 & 0.549 & 1.000 & 0.602 \\ 0.548 & 0.741 & 0.675 & 0.602 & 1.000 \end{pmatrix}$$

	AMZN	AAPL	GOOG	FB	MSFT
AMZN	1.000	0.578	0.493	0.509	0.548
AAPL	0.578	1.000	0.685	0.545	0.741
GOOG	0.493	0.685	1.000	0.549	0.675
FB	0.509	0.545	0.549	1.000	0.602
MSFT	0.548	0.741	0.675	0.602	1.000

[10]We take 178 returns backwards from December 2019 to construct R_{2019}.

FIGURE 1. Abnormal returns in 2020

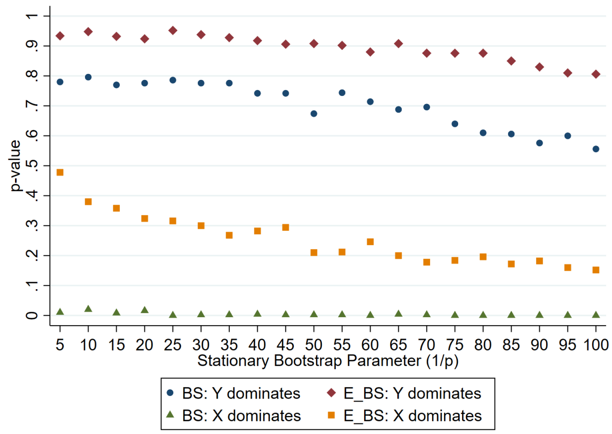


$$\hat{\sigma}_Y = \begin{pmatrix} 1.000 & 0.684 & 0.788 & 0.794 & 0.861 \\ 0.684 & 1.000 & 0.660 & 0.710 & 0.733 \\ 0.788 & 0.660 & 1.000 & 0.825 & 0.778 \\ 0.794 & 0.710 & 0.825 & 1.000 & 0.872 \\ 0.861 & 0.733 & 0.778 & 0.872 & 1.000 \end{pmatrix}.$$

We can observe that most of the off-diagonal entries have changed to more positive correlation. Thus, the values of sample correlations appear to be consistent with the null hypothesis of correlation dominance $\text{cov}(X_i, X_j) \leq \text{cov}(Y_i, Y_j)$ for all $1 \leq i \leq j \leq 5$. To verify this, we perform the correlation dominance test $C_T$and the supermodular dominance test $S_T$to $X = R_{2019}$ and $Y = R_{2020}$. As can be seen from Figure 2, we cannot reject the null of $\text{cov}(X_i, X_j) \leq \text{cov}(Y_i, Y_j)$ while we do reject the null of $\text{cov}(Y_i, Y_j) \leq \text{cov}(X_i, X_j)$ in most cases.

Since we use daily data, there may some concern about microstructure noise and its effects on our estimated covariance matrix. Therefore, we consider a multivariate version of the Roll

FIGURE 2. p-values of the correlation dominance test C_T



(1987) model that allows for bid ask bounce and use this to adjust the covariance matrix estimators. Suppose that

$$p_{it} = p_{it}^* + \frac{s_i}{2} Q_{it}, \quad p_{it}^* = p_{i,t-1}^* + \varepsilon_{it},$$

where $Q_{it}$is the trade direction indicator, which is assumed for now to be iid with +1 and -1 equally likely and s_i are the individual spreads. Then

$$\Delta p_{it} = \varepsilon_{it} + \frac{s_i}{2} \Delta Q_{it},$$

which implies that observed returns $r_{it} = \Delta p_{it}$are autocorrelated like an $MA(1)$process. We can identify $s_i, \sigma_{\varepsilon i}^2$from the variance and autocovariance of observed returns asset by asset as was done in the original Roll paper. We consider the multivariate extension of this where we suppose that for any assets $i, j$the order flow is potentially contemporaneously correlated so that

$$E\left(Q_{it} Q_{jt}\right) = \omega_{ij},$$

and we suppose that the efficient returns are themselves contemporaneously correlated with $E\left(\varepsilon_{it}\varepsilon_{jt}\right) = \sigma_{ij}$.Let $\Omega = (\omega_{ij})$and $\Sigma = (\sigma_{ij})$,where the univariate models imply that $\omega_{ii} = 1$.It follows that

$$E\left(r_t r_{t-j}^{\intercal}\right) = \begin{cases} \Sigma + \frac{1}{2}\left(ss^{\intercal} \odot \Omega\right) & \text{if } j = 0 \\ \frac{-1}{4}\left(ss^{\intercal} \odot \Omega\right) & \text{if } j = 1 \\ 0 & \text{if } j > 1, \end{cases},$$

where $\odot$ denotes the matrix Hadamard product. Therefore, it follows that

$$\Sigma = E(r_t r_t^{\intercal}) + E(r_t r_{t-1}^{\intercal}) + E(r_{t-1} r_t^{\intercal}). \tag{15}$$

The efficient covariance matrices of X and Y estimated based on (15) are as follows:

$$\hat{\Sigma}_X = \begin{pmatrix} 0.923 & 0.572 & 0.379 & 0.335 & 0.449 \\ 0.572 & 0.979 & 0.740 & 0.465 & 0.598 \\ 0.379 & 0.740 & 0.834 & 0.388 & 0.579 \\ 0.335 & 0.465 & 0.388 & 0.936 & 0.557 \\ 0.449 & 0.598 & 0.579 & 0.557 & 0.865 \end{pmatrix}, \ \hat{\Sigma}_Y = \begin{pmatrix} 0.409 & 0.174 & 0.229 & 0.145 & 0.127 \\ 0.174 & 0.637 & 0.189 & 0.302 & 0.198 \\ 0.229 & 0.189 & 0.555 & 0.263 & 0.117 \\ 0.145 & 0.302 & 0.263 & 0.285 & 0.083 \\ 0.127 & 0.198 & 0.117 & 0.083 & 0.082 \end{pmatrix}.$$

We can see that $\hat{\Sigma}_Y$ differs more from $\hat{\sigma}_Y$ than $\hat{\Sigma}_X$ does from $\hat{\sigma}_X$. We conduct the correlation dominance test (denoted E_BS) based on $\hat{\Sigma}_X$ and $\hat{\Sigma}_Y$ and Figure 2 shows the resulting p-values for various stationary bootstrap parameters.

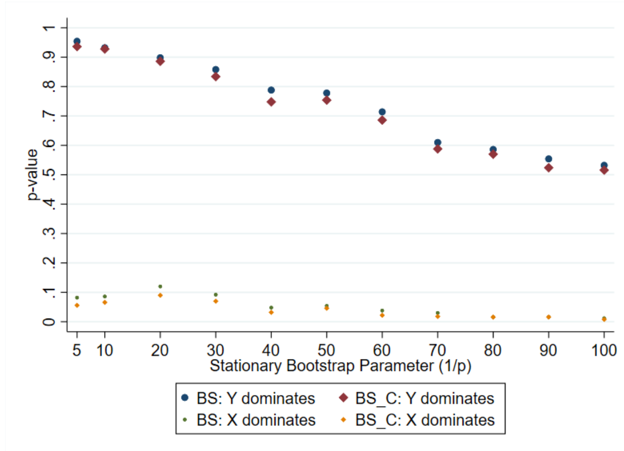
This phenomenon has been remarked upon before during other crisis periods, for example Morgan (2010) discussed the so-called correlation bubble that was experienced during the GFC. We have found evidence of a very strong effect whereby the pairwise correlations between these pivotal companies have uniformly increased between 2019 and 2020.

Given the correlation dominance results, we further test if there is a significant evidence of supermodular dominance. We used $a_0 = 0.9$ and the resulting p-values are plotted in Figure 3.[11] We find evidence in favor of $X \leq_{sm} Y$ because we can reject the null of $Y \leq_{sm} X$, while we do not reject the null of $X \leq_{sm} Y$ across different choices of the stationary bootstrap parameters.

6.2. **Knowledge spillover depending on distance.** In this subsection, we compare the effect of geographic and distant knowledge spillover, measured by patent citations, on social welfare. The issue of geographic knowledge spillover has received considerable attention in the literature; Black (2005) provides an extensive summary of the relationship between geography and innovation, see chapter 3 of this book for details. Some of the literature focusses on patent citations as a measure of knowledge spillover, including Jaffe et. al. (2000) who used the survey method to verify if patent citations can be an indicator of knowledge spillover. They

[11]The results are robust to the various choices of a_0 we considered.

FIGURE 3. p-values of the supermodular dominance test S_T



imply that the aggregate citation flows can serve as proxies for the intensity of knowledge spillover between groups of organizations or between countries. Jaffe et. al. (1993) also found evidence that in the U.S., citations to patents are geographically localized and these effects are quite large and quite statistically significant. They compare the frequency of localized citing between the original patents and the citing patents with that between the original patents and control group. Here, the control groups consist of patents with the same class and application year, and thus have the same temporal and technological distributions, with each citing patent. They found that citations are quantitatively and statistically significantly more localized than the controls.

While they approached this by manually matching a control group to the original patents, we believe that directly comparing localized and distant citing could provide a novel perspective to the literature. Therefore, we focus on comparing the effect of geographic and distant knowledge spillovers, measured by patent citations, on social welfare. Although the previous research matches temporal distributions of different patents only by their assigned dates, we directly utilize the overtime distributions of distant and localized citations on a monthly level, and test their equivalences by the Kolmogorov-Smirnov test. We also match the technological distributions of patents by comparing those with the same subclass, which is a four-digit classification of patent characteristics.

Theoretically, let f be a supermodular, monotone, and symmetric social welfare function. Then if either localized or distant citing supermodularly dominates the other, by the definition of supermodular dominance, it is expected to engender higher social welfare, i.e., $Ef(X) \leq Ef(Y)$holds for respective $X$and Y. We will test the hypotheses of both directions; one that tests the dominance of localized citing, and the other that tests the dominance of distant citing, to see if the result is reverted. If we reject in one case and fail to do so in another case, then we could conclude that there is indirect evidence of the supermodular dominance existing between the numbers of citations coming from within and outside of the state. In this case, by the definition of supermodular dominance, we can conjecture that $Ef(X) \leq Ef(Y)$or $Ef(Y) \leq Ef(X)$holds for supermodular, monotone, and symmetric social welfare function f, and thus, either localized or distant citing might engender greater expected social welfare than the other.

The data is obtained from the National Bureau of Economic Research (NBER) Patent Database. The NBER patent data comprise detail information on millions of U.S. patents, including patent number, date of assignment, assignee ID, country, state, and city of inventors, category[12] and subclass[13] of patents, citing patents, etc. There have been enormous efforts to increase the usability of this data set and among them we use Lee (2019) which provides information for patents assigned from 1976 to 2017. We drop the data with missing state and assignee ID information and confine our interest to the 50 American states.

Table 5 summarizes the descriptive statistics of the rank of the 50 states by the total number of patents, and the number of patents assigned in total and by patent category. Here, and throughout this paper, the location of a patent is identified as that of the first inventor of the patent. Out of the 50 American states, the top five states that have issued the most patents over the time period of the data are California (CA), New York (NY), Texas (TX), Massachusetts (MA), and New Jersey (NJ).

Table 6 reports the descriptive statistics of the rank of 50 states by the number of localized and distant citing. The third and the second columns represent the number of citations from within state and the rank of states by that, and the last two columns represent those of citations coming from different states. Overall, the two ranks seem to match closely. CA

[12]There are six different categories of patents: (1) chemical, (2) computers and communications, (3) drugs and medical, (4) electrical and electronics, (5) mechanicals, and (6) others.

[13]The subclass of a patent is defined as the first four letters of International Patent Classification (IPC).

Table 5. The number of patents assigned in the 50 U.S. states (1976-2017)

State	Rank	Number of patents assigned						
		in total	by category					
			1	2	3	4	5	6
CA	1	1,598,717	88,601	652,928	273,735	320,094	205,796	48,565
NY	2	522,921	71,844	150,219	54,419	120,551	102,138	19,918
TX	3	456,013	73,601	157,721	30,924	73,911	106,158	10,698
MA	4	336,669	29,664	95,146	88,890	55,751	52,655	12,506
NJ	5	327,075	67,480	84,292	78,332	37,613	43,479	13,325
MI	6	298,931	41,128	26,619	21,922	38,464	156,149	12,438
IL	7	295,160	47,353	61,659	35,008	39,416	87,898	21,534
PA	8	278,438	57,861	34,305	59,110	46,838	68,099	10,037
WA	9	261,464	11,728	150,752	26,207	30,233	33,908	7,301
OH	10	254,696	64,137	23,909	37,556	27,862	84,661	14,217
MN	11	225,106	29,741	51,472	54,722	28,905	48,349	10,618
FL	12	169,346	16,711	42,830	26,400	26,588	46,065	9,681
NC	13	153,366	14,246	54,494	21,803	22,532	31,673	7,766
CT	14	151,120	20,420	21,466	32,961	24,222	44,113	6,752
CO	15	132,055	10,612	49,846	19,086	22,451	24,217	5,031
WI	16	127,858	17,714	12,284	21,840	17,822	48,315	8,904
IN	17	114,555	14,738	9,449	30,013	17,561	35,751	6,064
AZ	18	114,428	7,558	39,455	8,724	29,870	24,342	3,849
OR	19	112,103	5,260	47,415	6,688	26,099	18,898	7,134
MD	20	108,792	12,046	23,323	31,286	17,986	19,448	3,812
GA	21	99,730	14,801	29,577	11,879	13,584	23,088	6,093
VA	22	92,326	11,509	28,217	8,871	15,560	21,646	5,886
MO	23	67,321	10,736	8,431	15,632	7,869	18,608	5,490
TN	24	57,304	13,375	4,247	12,058	7,010	16,963	3,281
UT	25	56,184	6,457	12,866	11,701	6,438	14,869	3,515
ID	26	55,349	2,630	10,039	1,126	31,092	8,746	1,240
IA	27	48,049	3,544	7,765	5,851	5,295	18,812	6,399
OK	28	44,478	14,785	2,754	2,465	3,951	18,751	1,348
NH	29	44,352	4,058	13,029	5,602	8,170	11,763	1,473
SC	30	42,372	8,869	2,812	3,104	5,341	19,009	2,937
KS	31	35,069	3,470	12,714	3,555	3,305	9,802	1,957
KY	32	35,026	4,927	8,262	3,628	3,503	12,503	1,930
DE	33	34,285	13,873	1,907	7,984	2,616	5,851	1,779
VT	34	32,421	1,352	7,771	1,033	17,320	3,838	894
LA	35	28,088	9,751	1,194	2,752	1,725	10,983	1,449
AL	36	26,536	4,229	4,653	4,242	3,953	8,048	1,215
NM	37	25,045	3,609	4,761	2,515	6,862	6,480	648
NV	38	24,260	1,638	5,365	1,477	2,574	4,289	8,787
RI	39	18,577	2,587	2,676	2,868	3,036	4,895	2,394
NE	40	14,191	1,259	2,388	1,856	2,222	4,912	1,460
WV	41	11,202	5,523	511	922	761	3,027	355
AR	42	10,034	1,413	1,040	1,215	1,249	4,130	926
ME	43	9,789	1,346	1,554	1,118	1,994	3,025	663
MS	44	9,191	1,581	909	1,389	1,114	3,117	1,019
MT	45	7,097	1,172	692	878	688	2,922	701
HI	46	5,029	449	1,049	930	622	1,424	535
ND	47	5,007	609	642	224	473	2,675	348
SD	48	4,513	415	655	499	415	1,957	544
WY	49	3,777	583	475	409	450	1,570	256
AK	50	2,302	300	166	229	190	1,198	190

Table 6. The number of citations patents received from within and outside of a state for the 50 U.S. states (1976-2017)

State	Rank_ss	Ncite_ss	Rank_ds	Ncite_ds
CA	1	742,164	1	497,126
OH	2	161,042	7	70,067
TX	3	135,210	2	208,397
MN	4	126,761	5	142,594
ID	5	86,765	9	51,987
NY	6	80,437	3	204,390
WA	7	80,309	4	154,191
MA	8	51,426	6	75,498
OR	9	39,912	8	69,347
NJ	10	31,493	10	48,066
NV	11	30,485	15	22,701
IL	12	30,224	13	30,697
MI	13	24,285	20	10,069
CT	14	21,295	11	40,017
OK	15	17,497	22	9,296
PA	16	12,652	17	14,626
NC	17	11,186	12	38,673
AZ	18	10,724	14	24,716
WI	19	10,121	21	9,730
IN	20	9,408	23	8,639
FL	21	8,176	16	15,421
CO	22	6,992	19	12,895
UT	23	6,299	24	8,000
GA	24	4,142	27	4,834
IA	25	3,902	29	3,845
NH	26	3,439	31	2,984
MD	27	3,213	26	4,918
VT	28	3,104	18	13,614
VA	29	2,358	30	3,782
TN	30	2,181	25	5,917
MO	31	1,767	28	4,208
DE	32	1,377	33	1,676
RI	33	1,281	34	1,415
KS	34	1,245	32	1,814
SC	35	1,183	35	694
KY	36	595	39	483
NM	37	571	40	352
MT	38	535	36	560
NE	39	511	38	486
LA	40	441	37	495
WV	41	235	43	141
AL	42	170	41	160
AR	43	118	45	61
WY	44	113	44	68
ME	45	95	42	147
ND	46	59	50	13
MS	47	44	46	48
HI	48	32	48	23
SD	49	27	47	29
AK	50	4	49	19

stands as the top state which have assigned patents that were cited most both from CA itself and from the other states.

Since a single firm can be located in more than one state and have issued patents of different subclasses, we take subgroups of firms based on the assignee ID, state, and subclass of patent, as we mentioned above. Firms are then selected based on this subgroup unit.[14] Focusing in on each state, we gather information on monthly numbers of citations firms received from the same state and from the other states. In specific, $Y = (Y_1, \ldots, Y_K)$is the collection of the numbers of citations that Kfirms received from patents assigned in the same state, while $X = (X_1, \ldots, X_K)$is the vector of mean citations that the same Kfirms received from the other states.[15] Note that here, $\mathbf{x}_k = (X_{k,1}, \ldots, X_{k,T})^\intercal$and $\mathbf{y}_k = (Y_{k,1}, \ldots, Y_{k,T})^\intercal$denote the vectors of numbers of citations of the same firm $k$for the months $t = 1, \ldots, 504$(i.e., from Jan. 1976 to Dec. 2017). We also confine our interest to those firms with at least 5 citations in total in order to focus on the effect of quality innovation.

For each state, after taking X and Y as described above for each firm, we apply our test on the numbers of citations, from the same state and the other states, that each firm received for its patents.[16] A grid is selected as follows: 0 to 50 by 1, 51 to 100 by 5, and by 10 after 101. We adopt a grid with denser grid points for lower values because in most cases, the number of citations a firm received in a month does not exceed a digit. We use stationary bootstrap critical values with $B = 500$bootstrap repetitions, $a_0 = 1.1$, and $1/p = 5$.[17]

The above procedure is carried out for the 50 US states. For most states with lower ranks in terms of the number of patents, we cannot perform the analysis because there is not much firm with patents that received at least five citations over the time period. Table 7 lists subclasses and their classifications for which firms satisfy the implications of supermodular dominance at least in one state. We observe several cases of supermodular dominance holding between distant and localized citing especially for subclasses that fall under the category of human

[14]For example, if firm i has issued patents of subclasses $s \in$\{AB12, CD34\} in states NY and OH, then the four possible subgroups \{i, AB12, NY\}, \{i, AB12, OH\}, \{i, CD34, NY\}, and \{i, CD34, NY\} are each treated as different entities, which are each referred to as a firm throughout this paper.

[15]The mean is taken over the number of states that have at least one record of citing the patents, for each firm.

[16]Due to the computational issue, we conduct our analysis on the subclasses which are assigned by no greater than 10 firms.

[17]The results are similar under different choices of the tuning parameters.

Table 7. Classification of subclass with potential supermodular dominance relationship between distant and localized citings

Classification	Subclass
A. Human necessities	A01H, A01K, A01N, A01N, A47B, A47L, A61B, A61C, A61F, A61K, A61L, A61M, A61N, A62B, A63B, A63F
B. Performing operations, Transporting	B01D, B01J, B01J, B05C, B05D, B08B, B22F, B23K, B23P, B24B, B25B, B26D, B27C, B29D, B31F, B32B, B32B, B41J, B60G, B60Q, B60R, B65D, B65H, B82B
C. Chemistry; Metallurgy	C01B, C03B, C07C, C07D, C07H, C07K, C08F, C08G, C08J, C08L, C08L, C09D, C09K, C10G, C10M, C11DC11D, C12N, C12P, C12Q, C25D, C25F
D. Textiles; Paper	D06L, D21F, D21H
E. Fixed constructions	E01F, E04C, E05B, E21B, E21B
F. Mechanical engineering; Lighting; Heating; Weapons; Blasting	F01B, F01N, F02G, F02K, F02M, F15C, F16H, F16J, F16K, F16L, F21V, F25B
G. Physics	G01B, G01K, G01L, G01N, G01R, G01S, G01V, G02B, G02C, G04C, G05B, G05F, G06F, G06K, G06Q, G06T, G07F, G08B, G09G, G10L, G11B, G11C
H. Electricity	H01C, H01G, H01H, H01J, H01L, H01M, H01Q, H01R, H02H, H02J, H03K, H03M, H04B, H04J, H04L, H04M, H04N, H04W, H05B, H05K

Source: World Intellectual Property Organization

necessities, performing operations and transporting, chemistry and metallurgy, physics, and electricity. Patents whose subclasses are classified as textiles and paper or fixed constructions constitute relatively limited cases of potential supermodular dominance.

Table 8 reports the test results where the implications of $Y \leq_{sm} X$holds and Table 9 reports those where the implications of $X \leq_{sm} Y$holds. Here, $Y \leq_{sm} X$means that we reject the null of Y dominates and accept the null of X dominates for all the bootstrap critical values, and similarly for $X \leq_{sm} Y$.The second columns of Table 8 and 9 are the rank of each state in terms of the total number of patents issued there. The results of the Kolmogorov-Smirnov tests are failure to reject the null of equal marginal distributions for most cases, and so is omitted.

It is notable that in most cases, distant citing supermodularly dominates localized citing, which means that citing between faraway states may lead to greater expected social welfare than localized citing. We observe some exceptions in 18 states but the numbers of subclasses

Table 8. Test result ($Y \leq_{sm} X$)

State	Rank	Number	Subclass
CA	1	30	A63B, B01D, B05D, B08B, B24B, B29D, B41J, C07C, C07H, C07K, C10G, C12P, C25F, E01F, G01B, G01S, G02C, G05F, G06T, G09G, G10L, H01C, H01J, H01M, H01Q, H01R, H03M, H04J, H04M, H05K
NY	2	21	B01D, B01J, B05D, B23K, B32B, B65H, C07C, C09D, C12Q, D21F, F16J, G01N, G01R, G01V, G02B, H01M, H01R, H03M, H04B, H04N, H05B
TX	3	19	A61B, A61K, A61N, A63B, B22F, C01B, C08F, C08L, C09K, F15C, G01N, G01V, G06K, G06Q, G11C, H01R, H03K, H04L, H05K
MA	4	13	A61M, B01D, G01B, G01N, G02B, G05B, G06K, G11B, H01L, H01M, H02J, H04B, H04W
NJ	5	15	A61C, A61F, B01J, C07D, C10M, C11D, C12Q, G02B, G06K, G10L, H01L, H04B, H04J, H04L, H04N
MI	6	9	A61B, B60Q, C07D, C08G, F01N, F02M, F16H, G02B, G06F
IL	7	11	A61B, A61K, A61M, B01D, B65D, C07D, G01N, H01L, H01R, H04L, H04M
PA	8	13	A01N, A61B, A61K, A61M, B25B, C07D, C08F, G02B, G06F, H01H, H01L, H01R, H02J
WA	9	12	A61K, B32B, C07D, G06K, G06Q, G10L, H01L, H01R, H04L, H04M, H04N, H04W
OH	10	7	A01N, A61B, B23K, D21H, G05B, G06F, G06Q
MN	11	11	A47L, A61K, B32B, F01B, F25B, G01L, G11B, H01G, H01L, H01M, H04L
FL	12	9	A61F, A61K, A61M, G02C, G06F, G08B, H01Q, H04B, H04L
NC	13	8	A61B, A61K, G02B, G06F, H01L, H01R, H04B, H04L
CT	14	11	A61B, A61K, A61L, A63F, C07D, F02K, G06F, G06Q, G07F, H01L, H02H
CO	15	4	A61B, A61K, G06Q, H04M
WI	16	6	A01H, B32B, C08J, D21F, G06F, H01M
IN	17	3	A61B, A61F, G01N,
AZ	18	5	A61B, G01B, G06F, H01L, H01M
OR	19	4	A63F, B41J, G01R, H04L,
MD	20	4	A61K, G06F, H04L, H04M,
GA	21	6	B60R, B65D, H02H, H04L, H04M, H04W,
VA	22	2	G06F, H04L
MO	23	2	A61B, G06F
TN	24	3	A61B, A61F, C08L
UT	25	4	A61M, B60R, E21B, G06F
IA	27	2	A01H, G08B
NH	29	3	B01J, C07C, E21B
NH	29	4	F02G, G06F, H01L, H01R
KS	31	4	G06F, H04J, H04M, H04W
DE	33	1	B32B
RI	37	1	A61B
NE	39	2	A63F, G07F

Table 9. Test result ($X \leq_{sm} Y$)

State	Rank	Number	Subclass
CA	1	2	B05C, F16K
NY	2	2	D21H, F16K
TX	3	3	A61M, B82B, G02B
MA	4	2	B05D, F16L
MI	6	1	E04C
IL	7	1	A63F
PA	8	2	F16L, G01K
OH	10	4	C03B, D06L, G05F, G06K
MN	11	1	A62B
FL	12	1	C07C
NC	13	2	A01K,B60G
CO	15	1	H05K
WI	16	1	G04C
IN	17	1	B23P
OR	19	6	B26D, C07H, C09D, C12N, C25D, G02B
MD	20	2	B25B, C11D
GA	21	2	B31F, E05B
VA	22	1	F21V
TN	24	1	B27C
UT	25	1	A47B
NM	35	1	H01L
VT	42	1	F02G

for such cases are minor compared to the other way around. The number of potential supermodular dominance relationships tend to increase as the rank of a state, in terms of the liveliness of patent activity, increases.

For patents issued in states and of subclasses not listed in Table 8 and 9, there is no evidence of supermodular dominance holding between citations from the same state and different states. In specific, they fall under either one of the following two cases: one where we reject the null of implication of supermodular dominance holding in both directions, and the other where we accept the null in both directions. The former case indicates there is no supermodular dominance between localized and distant citations whatsoever, and the latter case implies that there is not enough evidence to conclude so or the two citations may have an equal distribution. In these cases, we cannot draw on the welfare implication regarding the effect of localized and distant citing.

Overall, our result implies that there are some subclasses in certain states, especially in states where patent activities are lively, that exhibit the evidence of implications of supermodular dominance holding between localized and distant citing, while for most other subclasses in most states, we cannot find evidence that such relationship exists. Also, the effect of localized and distant citing on social welfare varies depending on patent locations and characteristics. Yet, it is worth noting that we are only testing necessary conditions of supermodular dominance. Analysis on a smaller than state-level could provide further information on the pattern of the dominance depending on the degree of localization.

7. Concluding Remarks

In this paper, we propose a consistent test of an implication of supermodular dominance, a notion that captures positive interdependence among inputs of supermodular functions. It can be utilized to see which input set creates greater utility, social welfare, production, etc. In the simulation, we found that the test has desirable asymptotic and finite sample properties under various parameter settings. We first applied our test to compare the interdependence of stock returns of major five firms before and after the COVID-19. We found the evidence that the returns after the COVID-19 supermodularly dominates that before the event, which suggests that the pandemic increased the interdependence between major stock returns. Another application of our test to comparing the effect on social welfare of geographic and distant knowledge spillover, measured by patent citation, indicates that firms having patents of certain subclasses in certain states, especially where patent activities are active, might have supermodular dominance holding between the localized and distant citing to their patents. In this case, we can conjecture that either localized or distant citing might engender greater expected social welfare than the other. For other cases, either there is no supermodular dominance relationship between localized and distant citing, or there is not adequate evidence to conclude so or the two may have an equal distribution.

References

[1] Amir, R. (2005). Supermodularity and Complementarity in Economics: An Elementary Survey. Southern Economic Journal , Jan., 2005, Vol. 71, No. 3 (Jan., 2005), pp. 636-660

[2] Andrews, D. W. K. and Pollard, D. (1994). An Introduction to Functional Central Limit Theorems for Dependent Stochastic Processes, *International Statistical Review* 62, 119-132.

[3] Athey, S. (2002). Monotone Comparative Statics Under Uncertainty, *The Quarterly Journal of Economics* 117(1), 187-223.

[4] Bäuerle, N. (1997). Inequalities for Stochastic Models via Supermodular Orderings, *Communications in Statistics, Stochastic Models* 13, 181-201.

[5] Black, G. (2005). *The Geography of Small Firm Innovation*, Kluwer Academic Publishers.

[6] Cooper, R. and Haltiwanger, J. (1996). Evidence on Macroeconomic Complementarities, *The Review of Economics and Statistics* 78(1), 78-93.

[7] Donald, S. G. and Hsu, Y. -C. (2016). Improving the Power of Tests of Stochastic Dominance, *Econometric Reviews* 35(4), 553-585.

[8] Goovaerts, M.J. and , J. Dhaene (1999). Supermodular ordering and stochastic annuities. Insurance: Mathematics and Economics 24 (1999) 281–290

[9] Jaffe, A., Trajtenberg, M. and Fogarty, M. (2000). Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors, *The American Economic Review* 90(2), 215-218.

[10] Han, H., Linton, O. B., Oka, T. and Whang, Y. -J., (2016). The Cross-Quantilogram: Measuring Quantile Dependence and Testing Directional Predictability Between Time Series, *Journal of Econometrics*, 193(1), 251-270.

[11] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Application*, Academic Press.

[12] Jaffe, A., Trajtenberg, M. and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, *The Quarterly Journal of Economics* 108, 577-598.

[13] Kizildemir, B. (2017). Supermodular comparison inequalities in option pricing and information inequalities. Doctoral thesis, Nanyang Technological University, Singapore.

[14] Kızıldemir, B. and N. Privault (2015). Supermodular ordering of Poisson arrays. Statistics and Probability Letters 98, 136-143.

[15] Meyer, M. and Strulovici, B. (2017). Beyond Correlation: Measuring Interdependence Through Complementarities, *Working Paper*.

[16] Müller, A. and Scarsini, M. (2000). Some Remarks on the Supermodular Order, *Journal of Multivariate Analysis* 73, 107-119.

[17] Morgan, J.P. (2010). Why We Have a Correlation Bubble. https://www.newconstructs.com/wp-content/uploads/2010/10/JP-Morgan-and-Correlation

[18] Lee, J. (2019). Korea Patent Data Project (KoPDP) `https://doi.org/10.7910/DVN/AUYERV`, Harvard Dataverse, V2, UNF:6:UYOuBGIK7GjpdWdAMYQX2A== [fileUNF].

[19] Linton, O. B., Maasoumi, E., and Whang, Y.-J. (2005). Consistent Testing for Stochastic Dominance Under General Sampling Schemes, *The Review of Economic Studies* 72(3), 735-765.

[20] Linton, O. B., Song, K. and Whang, Y.-J. (2010). An Improved Bootstrap Test of Stochastic Dominance, *Journal of Econometrics* 154, 186-202.

[21] Patton, A., Politis, D.N. and White, H. (2009). Correction to "Automatic Block-Length Selection for Dependent Bootstrap" by D. Politis and H. White, *Econometric Reviews* 28, 372-375.

[22] Politis, D.N. and Romano, J.P. (1994a). The Stationary Bootstrap, *Journal of American Statistical Association* 89, 1303-1313.

[23] Politis, D.N. and Romano, J.P. (1994b). Limit Theorem for Weakly Dependent Hilbert Space Valued Random Variables with Application to the Stationary Bootstrap, *Statistica Sinica* 4, 461-476.

[24] Politis, D.N. and White, H. (2004). Automatic Block-Length Selection for Dependent Bootstrap, *Econometric Reviews* 23, 53-70.

[25] Müller, A. and M. Scarsini (2000). Some Remarks on the Supermodular Order, *Journal of Multivariate Analysis* 73, 107-119.

[26] Rapuch, G., and T. Thierry Roncalli (2004). Some Remarks on Two - Asset Options Pricing and Stochastic Dependence of Asset Prices. Journal of Computational Finance, Vol. 7, No. 4,

[27] Topkis, D. (1979). Equilibrium Points in Nonzero-sum n-person Submodular Games, *Siam J. Control and Optimization* 17(6), 773-787.

[28] van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*, Springer.

[29] Vives, X. (1985). Nash Equilibrium in Oligopoly Games with Monotone Best Responses, *CARESS Working Paper*, No.85-10.

[30] Vives, X. (1990). Nash Equlibrium with Strategic Complementarities, *Journal of Mathematical Economics* 19, 305-321.

[31] Vives, X. (2005). Games with Strategic Complementarities: New Applications to Industrial Organization, *International Journal of Industrial Organization* 23, 625-637.

[32] World Intellectual Property Organization. (n.d.). IPC Publication. `https://www.wipo.int/classifications/ipc/ipcpub.`

Appendix

In Appendix, we provide proofs of Theorems in the main text.

Lemma A1. (i) *Suppose that Assumption 3.1 holds. Then, we have*

$$\nu_{\mathbf{i},T}(\cdot) := \sqrt{T}\left(\hat{D}_{\mathbf{i}}(\cdot) - D_{\mathbf{i}}(\cdot)\right) \Rightarrow \nu_{\mathbf{i}}(\cdot) \text{ in } l^{\infty}(\mathcal{X})$$

jointly for $\mathbf{i} \in \mathcal{I}$, *where* $\hat{D}_{\mathbf{i}}(\cdot) = (\hat{D}_{\mathbf{i}}^{U}(\cdot), \hat{D}_{\mathbf{i}}^{L}(\cdot))^{\intercal}$, $D_{\mathbf{i}}(\cdot) = (D_{\mathbf{i}}^{U}(\cdot), D_{\mathbf{i}}^{L}(\cdot))^{\intercal}$, *and* $\nu_{\mathbf{i}}(\cdot) = (\nu_{\mathbf{i}}^{U}(\cdot), \nu_{\mathbf{i}}^{L}(\cdot))^{\intercal}$ *is a mean zero Gaussian process with covariance kernel* $C(x_1, x_2) = \lim_{T\to\infty} E\nu_{\mathbf{i},T}(x_1)\nu_{\mathbf{i},T}(x_2)^{\intercal}$.

(ii) *Suppose that Assumptions 3.1 and 3.2 hold. Then, we have*

$$\sqrt{T}\left(\hat{D}_{\mathbf{i}}^{*}(\cdot) - \hat{D}_{\mathbf{i}}(\cdot)\right) \Rightarrow \nu_{\mathbf{i}}(\cdot) \text{ in } l^{\infty}(\mathcal{X})$$

conditional on $\mathbf{W}$ *in* P *jointly for* $\mathbf{i} \in \mathcal{I}$, *where* $\hat{D}_{\mathbf{i}}^{*}(\cdot) = (\hat{D}_{\mathbf{i}}^{U*}(\cdot), \hat{D}_{\mathbf{i}}^{L*}(\cdot))^{\intercal}$, *and* $\nu_{\mathbf{i}}(\cdot)$ *denotes a mean zero Gaussian process with covariance kernel* $C(x_1, x_2)$ *defined in (i).*

Proof of Lemma A1: (i) Let $\xi_{\mathbf{i},t}$for $\mathbf{i} \in \mathcal{I}$be a generic random variable that denotes $\max\{X_{i_1,t}, \ldots, X_{i_k,t}\}$, $\max\{Y_{i_1,t}, \ldots, Y_{i_k,t}\}$, $\min\{X_{i_1,t}, \ldots, X_{i_k,t}\}$or $\min\{Y_{i_1,t}, \ldots, Y_{i_k,t}\}$. Under Assumption 3.1, the class functions $\mathcal{F}_{\mathbf{i}} = \{1(\xi_{\mathbf{i},t} \leq x) : x \in \mathcal{X}\}$satisfies the $L_2(P)$-continuity condition

$$\sup_{t\geq 1} E \sup_{x'\in\mathcal{X}:|x-x'|\leq r} |1(\xi_{\mathbf{i},t} \leq x') - 1(\xi_{\mathbf{i},t} \leq x)|^2 \leq Cr$$

and hence its bracketing number satisfies $N_{[]}(\epsilon, \mathcal{F}_{\mathbf{i}}, L_2(P)) < C\epsilon^{-2}$. Also, let

$$\nu_{\mathbf{i},T}^{\xi}(\cdot) := \sqrt{T}(\mathbb{P} - P)f = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[1(\xi_{\mathbf{i},t} \leq \cdot) - E1(\xi_{\mathbf{i},t} \leq \cdot)\right],$$

denote the empirical process indexed by $f \in \mathcal{F}_{\mathbf{i}}$. Then, by the CLT for bounded strong mixing sequences (Hall and Heyde, 1980, Corollary 5.1) and Assumption 3.1(b), the finite dimensional distribution of $(\nu_{\mathbf{i},T}^{\xi}(x_1), \ldots, \nu_{\mathbf{i},T}^{\xi}(x_J))$converges to a normal distribution for all $(x_1, \ldots, x_J) \in \mathbb{R}^J$. Therefore, using the functional CLT of Andrews and Pollard (1994, Collorary 2.3) and the mixing condition in Assumption 3.1(a), we can show that the stochastic process $\nu_{\mathbf{i},T}^{\xi}(\cdot)$converges weakly to a Gaussian process $\nu_{\mathbf{i}}^{\xi}(\cdot)$with mean 0, covariance kernel $C_{\mathbf{i}}^{\xi}(x_1, x_2) = \lim_{T\to\infty} E\nu_{\mathbf{i},T}^{\xi}(x_1)\nu_{\mathbf{i},T}^{\xi}(x_2)$and continuous sample paths. The weak convergence holds jointly over $\mathbf{i} \in \mathcal{I}$and different choices of $\xi_{\mathbf{i},t}$, because the limiting Gaussian processes are separable and hence the joint convergence is equivalent to marginal convergence of each process (van der Vaart and Wellner (1996, Theorem 1.4.8)). This establishes Lemma A1(i).

(ii) The result of Lemma A1(ii) follows from the result of Lemma A1(i) and the bootstrap CLT of Politis and Romano (1994b, Theorem 3.1) because $\{Z_t(\cdot) := 1(\xi_{\mathbf{i},t} \leq \cdot) : t \geq 1\}$can be regarded as stationary Hilbert space valued random variables which are bounded and satisfy the mixing condition $\sum_{s=1}^{\infty} \alpha_Z(s) < \infty$under Assumption 3.1(a). □

Proof of Lemma 3.1: Let $\mathcal{B}_{\mathbf{i}} := \cup_{j=1}^{3} \mathcal{B}_{\mathbf{i},j}$. Write

$$\begin{aligned} S_T &= \sum_{\mathbf{i}\in\mathcal{I}} \left(T \int_{\mathcal{X}\setminus\mathcal{B}_{\mathbf{i}}} \left\{ [\hat{D}_{\mathbf{i}}^U(x)]_+^2 + [\hat{D}_{\mathbf{i}}^L(x)]_+^2 \right\} dx + T \int_{\mathcal{B}_{\mathbf{i}}} \left\{ [\hat{D}_{\mathbf{i}}^U(x)]_+^2 + [\hat{D}_{\mathbf{i}}^L(x)]_+^2 \right\} dx \right) \\ &=: (A_{1,T} + A_{2,T}). \end{aligned}$$

Note that whenever $x \in \mathcal{X} \setminus \mathcal{B}_{\mathbf{i}}$, we have $\max\{D_{\mathbf{i}}^U(x), D_{\mathbf{i}}^L(x)\} < -a_{T,1}$ under the null hypothesis. Therefore, for each $\mathbf{i} \in \mathcal{I}$, we have

$$\begin{aligned} A_{1,T} &\leq \sum_{\mathbf{i}\in\mathcal{I}} \int_{\mathcal{X}\setminus\mathcal{B}_{\mathbf{i}}} \left\{ [\sqrt{T}(\hat{D}_{\mathbf{i}}^U(x) - D_{\mathbf{i}}^U(x)) - \sqrt{T}a_{T,1}]_+^2 + [\sqrt{T}(\hat{D}_{\mathbf{i}}^L(x) - D_{\mathbf{i}}^L(x)) - \sqrt{T}a_{T,1}]_+^2 \right\} dx \\ &\leq C \sum_{\mathbf{i}\in\mathcal{I}} \left\{ \left[\sup_{x\in\mathcal{X}} \sqrt{T} \left| \hat{D}_{\mathbf{i}}^U(x) - D_{\mathbf{i}}^U(x) \right| - \sqrt{T}a_{T,1} \right]_+^2 + \left[\sup_{x\in\mathcal{X}} \sqrt{T} \left| \hat{D}_{\mathbf{i}}^L(x) - D_{\mathbf{i}}^L(x) \right| - \sqrt{T}a_{T,1} \right]_+^2 \right\}. \end{aligned} \tag{A.1}$$

This implies that $P(A_{1,T} = 0) \to 1$, because $\sup_{x\in\mathcal{X},\mathbf{i}\in\mathcal{I}} \sqrt{T} \left| \hat{D}_{\mathbf{i}}^U(x) - D_{\mathbf{i}}^U(x) \right| = O_p(1)$ by Lemma A1 and $\sqrt{T}a_{T,1} \to \infty$ by Assumption 3.3.

On the other hand,

$$\begin{aligned} A_{2,T} &\geq \sum_{\mathbf{i}\in\mathcal{I}} T \left\{ \int_{\mathcal{B}_{\mathbf{i},1}} [\hat{D}_{\mathbf{i}}^U(x)]_+^2 dx + T \int_{\mathcal{B}_{\mathbf{i},2}} [\hat{D}_{\mathbf{i}}^L(x)]_+^2 dx + T \int_{\mathcal{B}_{\mathbf{i},3}} \left\{ [\hat{D}_{\mathbf{i}}^U(x)]_+^2 + [\hat{D}_{\mathbf{i}}^L(x)]_+^2 \right\} dx \right\}. \\ &=: \bar{S}_T. \end{aligned} \tag{A.2}$$

Also, we have

$$A_{2,T} - \bar{S}_T = \sum_{\mathbf{i}\in\mathcal{I}} T \left\{ \int_{\mathcal{B}_{\mathbf{i},1}} [\hat{D}_{\mathbf{i}}^L(x)]_+^2 dx + \int_{\mathcal{B}_{\mathbf{i},2}} [\hat{D}_{\mathbf{i}}^U(x)]_+^2 dx \right\}$$
$$\leq C \sum_{\mathbf{i}\in\mathcal{I}} \left\{ \left[\sup_{x\in\mathcal{X}} \sqrt{T} \left| \hat{D}_{\mathbf{i}}^L(x) - D_{\mathbf{i}}^L(x) \right| - \sqrt{T} a_{T,2} \right]_+^2 + \left[\sup_{x\in\mathcal{X}} \sqrt{T} \left| \hat{D}_{\mathbf{i}}^U(x) - D_{\mathbf{i}}^U(x) \right| - \sqrt{T} a_{T,2} \right]_+^2 \right\}. \tag{A.3}$$

Now, by (A.2), (A.3) and the arguments similar to those for (A.1), we have $P(A_{2,T} = \bar{S}_T) \to 1$. This establishes Lemma 3.1, as desired. □

Proof of Theorem 3.2: Let

$$\tilde{S}_T := \sum_{\mathbf{i}\in\mathcal{I}} T \left\{ \int_{\mathcal{B}_{\mathbf{i},1}} [\hat{D}_{\mathbf{i}}^U(x) - D_{\mathbf{i}}^U(x)]_+^2 dx + \int_{\mathcal{B}_{\mathbf{i},2}} [\hat{D}_{\mathbf{i}}^L(x) - D_{\mathbf{i}}^L(x)]_+^2 dx \right\}$$
$$+ \sum_{\mathbf{i}\in\mathcal{I}} T \int_{\mathcal{B}_{\mathbf{i},3}} \left\{ [\hat{D}_{\mathbf{i}}^U(x) - D_{\mathbf{i}}^U(x)]_+^2 + [\hat{D}_{\mathbf{i}}^L(x) - D_{\mathbf{i}}^L(x)]_+^2 \right\} dx.$$

By Lemma 3.1, under the null hypothesis, we have

$$P\left(S_T \leq \tilde{S}_T\right) \to 1. \tag{A.4}$$

Let $\mathcal{B}_{\mathbf{i},j}^0$ denote $\mathcal{B}_{\mathbf{i},j}$ with $a_{T,1}$ replaced by $a_{T,2}$ for $j = 1, 2, 3$. We have

$$P\left(\mathcal{B}_{\mathbf{i},j} \leq \hat{\mathcal{B}}_{\mathbf{i},j} \leq \mathcal{B}_{\mathbf{i},j}^0\right) \to 1 \tag{A.5}$$

using the arguments of Linton, Song and Whang (2010, Proof of Theorem 2). Define $\tilde{S}_T^*$ to be the bootstrap statistic S_T^* (defined in (11)) with $\hat{\mathcal{B}}_{\mathbf{i},j}$ replaced by $\mathcal{B}_{\mathbf{i},j}$ for $j = 1, 2, 3$. Let $\tilde{c}_{T,\alpha}^*$ be the $(1-\alpha)$ quantile of the bootstrap distribution of $\tilde{S}_T^*$. Then, (A.5) implies that

$$P\left(c_{T,\alpha}^* \geq \tilde{c}_{T,\alpha}^*\right) \to 1. \tag{A.6}$$

There exists a subsequence $\{k_T : T \geq 1\} \subset \{T : T \geq 1\}$ such that

$$\limsup_{T\to\infty} P\left\{S_T > c_{T,\alpha,\eta}^*\right\} = \lim_{T\to\infty} P\left\{S_{k_T} > c_{k_T,\alpha,\eta}^*\right\}, \tag{A.7}$$

where S_{k_T} and $c_{k_T,\alpha,\eta}^*$ are the same as S_T and $c_{T,\alpha,\eta}^*$, except that the sample size T is replaced by k_T. Let $\tilde{\sigma}_T = Var(\tilde{S}_T)$. Since $\{\tilde{\sigma}_T : T \geq 1\}$ is a bounded non-negative sequence, there exists a further subsequence $\{l_T : T \geq 1\} \subset \{k_T : T \geq 1\}$ such that $\tilde{\sigma}_{l_T}$ converges. Consider first the case $\lim_{T\to\infty} \tilde{\sigma}_{l_T} > 0$. Then,

$$\begin{aligned} P\left\{S_{l_T} > c^*_{l_T,\alpha,\eta}\right\} &\leq P\left\{S_{l_T} > \tilde{c}^*_{l_T,\alpha}\right\} + o(1) \\ &\leq P\left\{\tilde{S}_{l_T} > \tilde{c}^*_{l_T,\alpha}\right\} + o(1) \qquad \text{(A.8)} \\ &\leq \alpha + o(1), \qquad \text{(A.9)} \end{aligned}$$

where the first inequality follows from the fact that $c^*_{T,\alpha,\eta} \geq c^*_{T,\alpha} \geq \tilde{c}^*_{T,\alpha}$ with probability approaching one by $(A.6)$, the second inequality follows from $(A.4)$, and the last inequality holds by the bootstrap consistency result in Lemma $A1$ and the uniform continuous mapping theorem (Linton, Song and Whang (2010, Lemma A1)).

Next, consider the other case $\lim_{T\to\infty} \tilde{\sigma}_{l_T} = 0$. We have

$$\begin{aligned} P\left\{S_{l_T} > c^*_{l_T,\alpha,\eta}\right\} &\leq P\left\{\tilde{S}_{l_T} > c^*_{l_T,\alpha,\eta}\right\} + o(1) \\ &\leq P\left\{\tilde{S}_{l_T} > \eta\right\} + o(1) \\ &\to 0, \qquad \text{(A.10)} \end{aligned}$$

where the first inequality follows from $(A.4)$, the second inequality holds by the definition of $c^*_{l_T,\alpha,\eta}$, and the last convergence to zero follows from the condition $\lim_{T\to\infty} \tilde{\sigma}_{l_T} = 0$ and the fact $\eta > 0$.

Now, (A.9) and (A.10) combine to yield

$$\limsup_{T\to\infty} P\left\{S_{l_T} > c^*_{l_T,\alpha,\eta}\right\} \leq \alpha. \qquad \text{(A.11)}$$

Since $P\left\{S_{k_T} > c^*_{k_T,\alpha,\eta}\right\}$ converges along $\{k_T\}$, it also converges along the subsequence $\{l_T\}$. Therefore, the result of Theorem 3.2(i) holds because the limsup in $(A.11)$ is equal to the limit in $(A.7)$.

(ii) Using an elementary inequality $[a+b]_+^2 \leq 2[a]_+^2 + 2[b]_+^2$, we have

$$
\begin{aligned}
S_T &= T \sum_{\mathbf{i} \in \mathcal{I}} \int_{\mathcal{X}} \{[\hat{D}_{\mathbf{i}}^U(x)]_+^2 + [\hat{D}_{\mathbf{i}}^L(x)]_+^2\} dx \\
&\geq \frac{1}{2} T \sum_{\mathbf{i} \in \mathcal{I}} \int_{\mathcal{X}} \{[D_{\mathbf{i}}^U(x)]_+^2 + [D_{\mathbf{i}}^L(x)]_+^2\} dx \\
&- T \sum_{\mathbf{i} \in \mathcal{I}} \int_{\mathcal{X}} \{[\hat{D}_{\mathbf{i}}^U(x) - D_{\mathbf{i}}^U(x)]_-^2 + [\hat{D}_{\mathbf{i}}^L(x) - D_{\mathbf{i}}^L(x)]_-^2\} dx. \qquad \text{(A.12)}
\end{aligned}
$$

The last term of (A.12) is $O_p(1)$ by the weak convergence result in Lemma A1(i). Also, under the alternative hypothesis H_1, we have

$$
T \sum_{\mathbf{i} \in \mathcal{I}} \int_{\mathcal{X}} \{[D_{\mathbf{i}}^U(x)]_+^2 + [D_{\mathbf{i}}^L(x)]_+^2\} dx \to \infty.
$$

The result of Theorem 3.2(ii) now holds because $c_{T,\alpha,\eta}^* = O_p(1)$ by Lemma A1(ii). □

Department of Economics, Seoul National University

Email address: dbchung@snu.ac.kr

Department of Economics, University of Cambridge

Email address: obl20@cam.ac.uk

Department of Economics, Seoul National University

Email address: whang@snu.ac.kr