**Economics Division
University of Southampton
Southampton SO17 1BJ, UK**

**Discussion Papers in
Economics and Econometrics**

Title: Conditional Stochastic Dominance Tests in Dynamic Settings

By Jesus Gonzalo and Jose Olmo

No.
1311

This paper has been accepted for publication in the International Economic Review.

Conditional Stochastic Dominance Tests in Dynamic Settings*

By Jesus Gonzalo[1] and Jose Olmo

This paper proposes nonparametric consistent tests of conditional stochastic dominance of arbitrary order in a dynamic setting. The novelty of these tests lies in the nonparametric manner of incorporating the information set. The test allows for general forms of unknown serial and mutual dependence between random variables, and has an asymptotic distribution that can be easily approximated by simulation. This method has good finite-sample performance. These tests are applied to determine investment efficiency between US industry portfolios conditional on the dynamics of the market portfolio. The empirical analysis suggests that Telecommunications dominates the other sectoral portfolios under risk aversion.

JEL classification: C1, C2, G1.

Keywords: Hypothesis testing, kernel estimation, lower partial moments, nonparametric regression, p-value transformation, stochastic dominance

1 Introduction

During the last thirty years, the interest in comparisons of random variables has shifted from hypothesis tests for the first and second statistical moments to more complex tests that consider the entire distribution of the data. The reason for this is twofold. First, the common belief is that the underlying generating processes are nonlinear and cannot be described by simple models of mean and variance. Second, the development of sophisticated mathematical and statistical techniques is based on empirical processes that allow for a comparison between distribution functions and higher statistical moments. The interest in testing for stochastic dominance between random variables has arisen in different theoretical and applied fields within statistics, economics and recently, finance. The comparison of wealth distributions between economies has been widely investigated in the literature (see McFadden (1989), Larsen and Resnick (1993), Kaur, Prakasa Rao and Singh

*Manuscript received October 2010; revised March 2012

[1]We acknowledge helpful discussions with Miguel Delgado, Juan Carlos Escanciano, Clive Granger, Oliver Linton, Oscar Martinez, Gabriel Montes-Rojas and Carlos Velasco. We are also grateful to seminar participants of ECARES (ULB) and to participants of North American Winter Meetings of the Econometric Society, Workshop in Computational and Financial Econometrics in Neuchatel, Far East and South Asia Meetings of the Econometric Society in Singapore, XXXIII Simposium of the Spanish Economic Association in Zaragoza, 4^{th} Tinbergen Institute Conference in Rotterdam, Royal Economic Society in University of Surrey and SOFIE 2009 European conference in Geneva. Financial support from the MICINN ECO2010-19357, ECO2011-22650, NEINVECON 06-11, EXCELECOM-CM grant and Bank of Spain (grant ER program) is gratefully acknowledged.

(1994), Anderson (1996), Davidson and Duclos (2000) and Barrett and Donald (2003), amongst others). The close relationship between the concept of stochastic dominance and expected utility maximization for rational investors has also produced a fertile area of research in finance (see Stone (1973), Porter (1974) or Fishburn (1977)). These authors discuss the link between stochastic dominance and portfolio efficiency. More recently, Shalit and Yitzhaki (1994) and Linton, Maasoumi and Whang (2005, LMW hereafter) extend this relationship to conditional portfolio efficiency and conditional stochastic dominance.

The concept of conditional stochastic dominance has been subject to different interpretations. Thus, Shalit and Yitzhaki (1994) define marginal conditional stochastic dominance as the probabilistic conditions under which all risk-averse individuals, conditional on a portfolio of assets, prefer to increase the share of a risky asset to that of another asset in the same portfolio. These authors study the implications of this definition in the efficiency of the market portfolio. LMW, however, econometrically analyze the implications of extending stochastic dominance and portfolio efficiency to a conditional, potentially dynamic, setting. These authors allow for serial and cross dependence between investment portfolios and develop hypothesis tests for conditional stochastic dominance with the aim of uncovering stochastically maximal investment strategies conditional on other explanatory factors. Related tests for stochastic dominance and portfolio efficiency are found in Post (2003), Kopa and Post (2009) and Scaillet and Topaloglou (2010), among others.

The statistical methods necessary to test for stochastic dominance of an arbitrary order are based on empirical processes and complex asymptotic theory. A seminal contribution is that of Barrett and Donald (2003) who develops tests for stochastic dominance between independent random variables in an independent and identically distributed (*iid*) framework. The asymptotic distribution of their family of test statistics follows a Gaussian process with a covariance function that depends on functions of the cumulative marginal distributions of the random variables, and hence cannot be tabulated. These authors propose a bootstrap procedure and a simulation method based on Hansen's (1996) p-value transformation to approximate the asymptotic distribution of the test. Their method also allows for different sample sizes for each random variable. The limitations of this method for the analysis of time series, which are used in most financial applications, are obvious and lead LMW to extend the method to propose consistent tests of stochastic dominance under general sampling schemes that include serial and mutual dependence between random variables. These authors work in a parametric framework in which the response variable is a function of sets of explanatory variables that can contain lags of the endogenous variable. Their method also permits working with the residuals of parametric models, and, therefore, developing tests of conditional stochastic dominance. Unfortunately, the estimation of model parameters invalidates the asymptotic theory developed in Barrett and Donald (2003) due to an extra term produced by estimation uncertainty that remains in the asymptotic distribution of the test. LMW solve this problem by implementing subsampling methods to approximate this distribution. This resampling method produces consistent estimates of the critical values of the test not only under the least favorable case given by the equality of functions but also on the boundary of the null hypothesis. The formulation of these authors is very flexible and allows for general conditioning schemes. The parametric nature of the method, potentially affected by model misspecification, and the choice of block size in the subsampling approximation of the critical values of the test are subject to criticism and discussion.

More recently, Linton, Song and Whang (2010) propose bootstrap tests that refine the method in LMW by achieving asymptotic sizes less than or equal to the nominal level uniformly over the probabilities in the null hypothesis. These tests lead to an improvement in the power over LMW but face the same potential problems discussed above. Delgado and Escanciano (2013) also propose bootstrap-based stochastic dominance tests with asymptotic sizes equal to the nominal level uniformly over the boundary points of the null hypothesis. In contrast to Linton, Song and Whang (2010), these authors focus on testing first-order stochastic dominance between nonparametric conditional distributions of *iid* random variables. Although this test can be easily extended to higher orders of stochastic dominance, the extension of the test to a dynamic time series framework appears more cumbersome.

The main contribution of this paper is to develop hypothesis tests of stochastic dominance of arbitrary orders under general conditioning schemes that, unlike LMW, do not require parametric specifications of the conditional dynamics. By a transformation of the different statistical moments of the random variables in terms of lower partial moments and the use of nonparametric kernel methods for stationary $\beta-$mixing processes, we can characterize the null hypothesis of stochastic dominance of an arbitrary order as a nonparametric regression between the difference of weighted functions of the random variables under comparison and the information set, approximated by a finite vector of regressors. This methodology is very flexible; estimators of the different quantities are obtained from standard nonparametric kernel regression methods, and the asymptotic theory follows from combining well-known results in nonparametric econometrics for conditional density estimation and regression. Our tests allow for general forms of serial and mutual dependence between the variables to be compared as well as those contained in the information set. The asymptotic distribution of the tests depends on nuisance parameters and hence, it cannot be tabulated; to overcome this issue we propose simulation methods that approximate the p-values of the tests. In particular, we discuss a multiplier method similar in spirit to the simulation method proposed in Hansen (1996) and more recently in Chernozhukov, Lee and Rosen (2012). The method is shown to work well for small sample sizes and for arbitrary orders of stochastic dominance.

This theory is applied to determine the efficiency of ten portfolios representing US industrial sectors: Nondurables, Durables, Manufactures, Energy, High Technology, Telecommunications, Shops, Health, Utilities and Others, conditional on the performance of a value-weighted market portfolio, spanning the period 1960-2009. Our results show that the Telecommunications sector dominates the High-Tech and Shop sectors for the second and third orders of stochastic dominance. Furthermore, at the 20% significance level, this portfolio also dominates for second and third orders of stochastic dominance the other sectoral portfolios.

The paper is structured as follows. Section 2 introduces the definition of stochastic dominance under general conditioning schemes and proposes hypothesis tests for stochastic dominance of arbitrary orders. Section 3 derives the asymptotic theory for these tests and discusses a simulation method to consistently approximate the asymptotic p-value of the test. In Section 4 we perform a Monte Carlo simulation experiment to study the finite sample performance of the proposed tests. Section 5 applies this testing method to assess stochastic dominance between US industrial sectors conditional on the dynamics of the market portfolio. Section 6 concludes; proofs are presented in a mathematical appendix.

2 Conditional Stochastic Dominance in Dynamic Models

This section extends the definition of stochastic dominance to general conditioning schemes and proposes consistent hypothesis tests for this condition based on nonparametric methods. Let $(Y_t^A, X_t)_{t\in\mathbb{Z}}$ and $(Y_t^B, X_t)_{t\in\mathbb{Z}}$ be two different $\mathbb{R}^{1+k}$ strictly stationary multivariate time series processes with an information set $I_t = \{(Y_{s-1}^A, Y_{s-1}^B, X_s), t-m+1 \leq s \leq t\}$ defined on a compact set $\Omega' \subset \mathbb{R}^q$ with $q=(k+2)m$. Let $F(y)$ be the marginal cumulative distribution function (cdf) corresponding to Y_t, $F_{I_t}(y) = P\{Y_t \leq y | I_t\}$ the corresponding distribution function conditional on the set I_t, and $f(\cdot)$ and $f_{I_t}(\cdot)$ the respective density functions. The marginal distribution and density functions of I_t are $F^{I_t}(\cdot)$ and $f^{I_t}(\cdot)$, respectively. The subindex s in m_s and m_{ss} denotes the first and second derivatives of a generic function $m(I_t)$ with respect to the component s of the multivariate vector I_t. The indexes A and B denote the random variables Y_t^A and Y_t^B that are defined on a compact set $\Omega \subset \mathbb{R}$; $(Y_t, I_t) \in \widetilde{\Omega} = \Omega \times \Omega'$.

The definition of unconditional $\gamma-$stochastic dominance of Y_t^B by Y_t^A for $1 \leq \gamma < \infty$ is

$$(1) \qquad \Psi_\gamma^A(y) \leq \Psi_\gamma^B(y), \text{ for all } y \in \Omega \subset \mathbb{R},$$

with strict inequality for some y (see Levy (2006)); $\Psi_\gamma(y) = \int_{-\infty}^{y} \Psi_{\gamma-1}(\tau)d\tau$ with $\Psi_1(y) = F(y)$. Using integration by parts for $\Psi_\gamma(y)$, we observe that the above definition can be expressed as

$$\int_{-\infty}^{y} (y-\tau)^{\gamma-1} f^A(\tau)d\tau \leq \int_{-\infty}^{y} (y-\tau)^{\gamma-1} f^B(\tau)d\tau \text{ for all } y \in \Omega \subset \mathbb{R}.$$

This characterization of stochastic dominance has been thoroughly discussed in early studies on portfolio efficiency (see Stone (1973), Porter (1974) or Fishburn (1977)). For the study of conditional stochastic dominance, we modify these definitions to incorporate the conditional distribution $F_{I_t}(\cdot)$.

Definition: Y_t^A *γ-stochastic dominates* Y_t^B *conditional on* I_t *for all* $t \in \mathbb{Z}$, *if and only if*

$$\Psi^A_{I_t,\gamma}(y) \leq \Psi^B_{I_t,\gamma}(y) \text{ for all } y \in \Omega \text{ and } t \in \mathbb{Z}, \tag{2}$$

where $\Psi_{I_t,\gamma}(y) = \int_{-\infty}^{y} \Psi_{I_t,\gamma-1}(\tau)d\tau$ *and* $\Psi_{I_t,1}(y) = F_{I_t}(y)$.

Using integration by parts, $\Psi_{I_t,\gamma}(y) = \int_{-\infty}^{y} (y-\tau)^{\gamma-1} f_{I_t}(\tau)d\tau$ and condition (2) reads as

$$\int_{-\infty}^{y} (y-\tau)^{\gamma-1} f^A_{I_t}(\tau)d\tau \leq \int_{-\infty}^{y} (y-\tau)^{\gamma-1} f^B_{I_t}(\tau)d\tau \text{ for all } y \in \Omega \text{ and } t \in \mathbb{Z}. \tag{3}$$

An alternative characterization of stochastic dominance is provided in terms of the class of all von Neumann-Morgenstern type utility functions; (see Lemma 1 in Fishburn (1977), Shalit and Yitzhaki (p. 671, 1994) for second stochastic dominance, or Definition 2 in LMW). The extension of these results to a conditional dynamic setting is straightforward and omitted for the sake of brevity. The difference from the unconditional (static) approach is that by testing dynamically for the stochastic dominance of one investment strategy over another, we can assess the optimality of the investor's decision as the information set varies over time.

Klecan, McFadden and McFadden (1991), Anderson (1996), Davidson and Duclos (2000) and more recently Barrett and Donald (2003), pioneered the development of hypotheses for arbitrary orders of stochastic dominance in an *iid* setting. The test is defined as

$$\sup_{y \in \Omega} \left(\Psi^A_\gamma(y) - \Psi^B_\gamma(y) \right) \leq 0.$$

The stationary version of this test under the presence of serial dependence in the data is developed in Scaillet and Topaloglou (2010). LMW, however, focus on dynamic tests of conditional stochastic dominance based on the analysis of residuals of time series regression models. The definition of conditional stochastic dominance in (2) and the characterization in (3) allow us to propose the following composite test for the hypothesis of conditional stochastic dominance in dynamic settings:

$$(4) \qquad H_{0,\gamma} : E[d_{t,\gamma}(y)|I_t] \leq 0 \text{ for all } y \in \Omega \text{ and } t \in \mathbb{Z},$$

with $d_{t,\gamma}(y) = (y - Y_t^A)^{\gamma-1} 1(Y_t^A \leq y) - (y - Y_t^B)^{\gamma-1} 1(Y_t^B \leq y)$, vs.

$$H_{1,\gamma} : E[d_{t,\gamma}(y)|I_t] > 0 \text{ for some } y \in \Omega \text{ and } t \in \mathbb{Z}.$$

The stationarity of the multivariate distribution of the random variables (Y_t^A, I_t) and (Y_t^B, I_t) implies that this condition can be expressed in terms of (Y_1^A, I_1) and (Y_1^B, I_1) as[2]

$$(5) \qquad H_{0,\gamma} : E[d_{1,\gamma}(y)|I_1 = x] \leq 0 \text{ for all } z = (x, y) \in \widetilde{\Omega}, \text{ vs.}$$

$$H_{1,\gamma} : E[d_{1,\gamma}(y)|I_1 = x] > 0 \text{ for some } z = (x, y) \in \widetilde{\Omega}.$$

The null hypothesis of these tests is composite, meaning that there are infinitely many conditions to be tested. Therefore, it is not clear in principle how one should derive the sampling distribution under the null hypothesis. Barrett and Donald (2003) and previous authors focus on the least favorable case under the null hypothesis. The advantage of this approach resides in its simplicity when deriving the asymptotic theory of the test. However, the use of the least favorable case as a null hypothesis results in the largest critical values possible. Romano and Wolf (2011), in a similar setting, also advocate this approach as a conservative but useful method to obtain critical values under composite null hypotheses. In our framework the least favorable case is given by $\widetilde{H}_{0,\gamma} : E[d_{1,\gamma}(y)|I_1 = x] = 0$ for all $z \in \widetilde{\Omega}$. The rejection of the null hypothesis $H_{0,\gamma}$ implies that Y_t^A does not dominate Y_t^B stochastically for order γ. The failure to reject this null hypothesis is a necessary condition for the presence of stochastic dominance of Y_t^A over Y_t^B. However, this test needs to be complemented with the reverse test characterized by swapping the roles of the random variables under both hypotheses ($H_{0,\gamma}^r : E[-d_{1,\gamma}(y)|I_1 = x] \leq 0$ for all $z \in \widetilde{\Omega}$ and $H_{1,\gamma}^r : E[-d_{1,\gamma}(y)|I_1 = x] > 0$ for some $z \in \widetilde{\Omega}$). The rejection of $H_{0,\gamma}^r$ against $H_{1,\gamma}^r$ implies that Y_t^A dominates Y_t^B stochastically; otherwise the hypothesis of equality of the quantities $\Psi_{I_1,\gamma}^A$ and $\Psi_{I_1,\gamma}^B$ cannot be rejected. Finally, if the null hypothesis $H_{0,\gamma}$ is rejected against $H_{1,\gamma}$ and $H_{0,\gamma}^r$ is rejected against $H_{1,\gamma}^r$, there is statistical evidence to claim that Y_t^A and Y_t^B are

[2]See Delgado and Escanciano (2007) for similar uses of this notation.

stochastically efficient (no dominance in either direction) with order γ.

3 Asymptotic Theory and P-value Approximation

This section introduces a family of test statistics for $H_{0,\gamma}$ and develops the corresponding asymptotic theory based on $\widetilde{H}_{0,\gamma}$. The asymptotic distribution of the tests is nonstandard and cannot be tabulated. To solve this problem, we also develop a simulation method that approximates the asymptotic p-values of the tests. The methodology is nonparametric and based on kernel estimators of the relevant quantities necessary for our study.

3.1 Asymptotic Theory

We consider the following nonparametric estimator of $\Psi_{I_t,\gamma}(y)$ for $I_t = x$, a fixed vector of dimension q:

$$\text{(6)} \qquad \widehat{\Psi}_{x,\gamma}(y) = \frac{n^{-1}\sum_{t=1}^{n}(y - Y_t)^{\gamma-1}1(Y_t \leq y)W_h\left(\frac{I_t - x}{h}\right)}{\widehat{f}^{I_1}(x)},$$

where $\widehat{f}^{I_1}(x) = n^{-1}\sum_{t=1}^{n} W_h\left(\frac{I_t-x}{h}\right)$ is the kernel estimator of the multivariate density function $f^{I_1}(x)$; and $W_h\left(\frac{I_t-x}{h}\right) = \prod_{s=1}^{q} h_s^{-1} w\left(\frac{I_{t,s}-x_s}{h_s}\right)$ where $w(\cdot)$ is a univariate kernel function. Note that $I_{t,s}$ and x_s denote the $s^{th}-$component of the multivariate random vectors I_t and x, respectively; h_s is the bandwidth parameter corresponding to the variable $I_{t,s}$.

Let $D_{n,\gamma}(z) = \widehat{\Psi}^A_{x,\gamma}(y) - \widehat{\Psi}^B_{x,\gamma}(y)$ be the empirical version of $\Psi^A_{x,\gamma}(y) - \Psi^B_{x,\gamma}(y)$ for the vector $z = (x, y) \in \widetilde{\Omega}$. This estimator, also expressed as

$$\text{(7)} \qquad D_{n,\gamma}(z) = \frac{n^{-1}\sum_{t=1}^{n} d_{t,\gamma}(y)W_h\left(\frac{I_t-x}{h}\right)}{\widehat{f}^{I_1}(x)},$$

can be interpreted as the Nadaraya-Watson kernel estimator (see Nadaraya (1965) and Watson (1964)) in the following dynamic stationary regression:

$$\text{(8)} \qquad d_{t,\gamma}(y) = g_\gamma(I_t, y) + u_t(y),$$

where $g_\gamma(I_t, y)$ is an unknown smooth function that depends on the value x that the information set I_t takes at time t. In the standard mean square error sense the function $g_\gamma(z)$ obtained from $I_t = x$ is interpreted as the conditional mean of $d_{t,\gamma}(y)$ given $I_t = x$, *i.e.* $g_\gamma(z) = E[d_{t,\gamma}(y)|I_t = x]$ with $z = (x, y)$, and $u_t(y)$ is the error term of the regression that satisfies $E[u_t(y)|I_t = x] = 0$. We further assume that the error process is *iid* for all $y \in \Omega$. This representation of the problem in terms of nonparametric mean regression allows us to write the null hypothesis in (5) as $H_{0,\gamma} : \sup_{z\in\widetilde{\Omega}} g_\gamma(z) \leq 0$ and the least favorable case $\widetilde{H}_{0,\gamma}$ as $g_\gamma(z) = 0$ for all $z \in \widetilde{\Omega}$. The asymptotic theory of the test exploits these characterizations of the null hypothesis.

In both theoretical and practical settings, nonparametric kernel estimation has been established as being relatively insensitive to the choice of the kernel function. The same cannot be said for bandwidth selection. The interpretation of (7) in terms of a nonparametric regression allows us to use standard least squares cross-validation methods to determine the optimal vector of bandwidth parameters. The advantage of this method over other alternatives, such as a rule of thumb or plug-in methods is that cross-validation automatically discards irrelevant information from the vector I_t (see Li and Racine (2007, p. 69)). We should acknowledge, however, that in practice, the use of nonparametric regression techniques can be challenging if the conditioning sets are defined by a large number of covariates. Unfortunately, there is no easy solution to this problem that is intrinsic to the nonparametric regression literature. Partial solutions to mitigate the problem involve imposing some structure on the nonparametric regression, as for example an additive model. It is well known that for kernel-based methods, two approaches are commonly used for estimating an additive model: the backfitting method (see Buja, Hastie and Tibshirani (1989) and Hastie and Tibshirani (1989)) and the marginal integration method proposed by Linton and Nielsen (1995) and Newey (1994), among other authors. We believe that the implementation of these methods is beyond the scope of this paper.

The interpretation of the test given by (8) also allows us to apply standard asymptotic theory on non-parametric regression models for weakly dependent data. We first require the following assumptions:

Assumptions:

A.1: The process $\{(Y_t^A, Y_t^B, I_t), t \in \mathbb{Z}\}$ is strictly stationary and β-mixing with β-mixing coefficients that satisfy $\beta(j) \leq C \exp(-C_1 j)$, with $C, C_1 > 0$ being constants. The β-mixing coefficient is defined as $\beta(j) = \sup_i E\left[\sup_{V\in\Im_{i+j}^n} \left\{P(V|\Im_1^i) - P(V)\right\}\right]$.

A.2: Let $f_{I_i|\Im_j^{i-1}}$ be the density of the conditional distribution $P\{I_i \leq x \mid I_j, \ldots, I_{i-1}\}$. There exist constants

$C_2, C_3 > 0$ such that

$$\sup_{i>\eta+1}\left\{P\left(\sup_{x\in\Omega}\left[|f_{I_i|\Im_1^{i-1}}(x)-f_{I_i|\Im_{i-\eta}^{i-1}}(x)|\right]>C\exp(-C_2\eta)\right)\right\}\leq C\exp(-C_3\eta)\ \forall\eta\geq 1$$

and

$$\sup_{i>1}\sup_{x\in\Omega}\left\{f_{I_i|\Im_1^{i-1}}(x)\right\}\leq C.$$

A.3: The joint *cdfs* of (I_1, Y_1^A) and (I_1, Y_1^B) are uniformly continuous on $\mathbb{R}^{q+1}$. The functions $f^{I_1}(x)$ and $g_\gamma(z)$ are three-times differentiable with respect to I_1, with derivative functions that satisfy the Lipschitz condition $|m(u)-m(v)|\leq C|u-v|$ for some $C>0$. The function $f^{I_1}(x)$ is bounded away from zero for $x\in\Omega'$.

A.4: The kernel function $w(\cdot)$ implicit in (6) is a symmetric, bounded on $[-1,1]$, and compactly supported probability distribution function. Defining $H_l(v)=|v|^l W(v)$ with $W(v)=\prod_{j=1}^{q} w_j(v)$, we assume that $|H_l(v)-H_l(u)|\leq C_2|u-v|$ for all $0\leq l\leq 3$ and some constant $C_2>0$.

A.5: Assume for simplicity that $h_s=h$ for $s=1,\ldots,q$. Then, as $n\to\infty$, $h\to 0$, $(nh^q)^{1/2}/log\ n\to\infty$, $log\ n/(n^{1/(q+1)}h)\to 0$ and $h^{q/2}log\ n\to 0$.

A.6: The conditional distributions of Y_t^A and Y_t^B given $\Im_t$ depend only on I_t, with $\Im_t$ the $\sigma-$field generated by the information set up to time t.

A.7: The sequence $u_t(y)$ is an *iid* process and satisfies that $E[u_t(y)\ |I_t=x]=0$ for all $z\in\widetilde{\Omega}$. This process is uniformly continuous on $y\in\Omega$ and $E[u_t^2(y)\ |I_t=x]$ is Lipschitz continuous and bounded away from zero on their support.

Assumptions A.1 and A.2 limit the extent of short weak dependence and allow us to apply the strong approximation results for density estimators with weakly dependent observations by density estimators from associated *iid* processes in Neumann (1998). A.7 assumes that the process $g_\gamma(I_t,y)$ not only captures conditional dependence in the mean but also the extant serial dependence in the time series $d_{t,\gamma}(y)$, for all $y\in\Omega$. The *iid* property imposed on $u_t(y)$ is need to apply the strong approximation results discussed in Theorem 1 and thereafter. We must acknowledge that this assumption is quite restrictive in a dynamic setting and extensions considering $u_t(y)$ to be a martingale difference sequence are very desirable. In fact, the martingale difference assumption is sufficient to obtain pointwise consistency and inference results. The rest of assumptions are standard for estimation and inference in nonparametric kernel methods. Under these assumptions we can apply the results on uniform convergence for nonparametric kernel regression estimators by Masry

(1996) and Hansen (2008) to our setting. In particular we have that

$$\sup_{z\in\widetilde{\Omega}}\left|D_{n,\gamma}(z)-g_{\gamma}(z)\right| = O\left(\sum_{s=1}^{q}h_s^2+\left(\frac{log\ n}{nh^q}\right)^{1/2}\right) \text{ as } n\to\infty$$

almost surely. This result extends the pointwise convergence in probability and shows that $D_{n,\gamma}(z)$ can be used for testing the composite hypothesis $H_{0,\gamma}$. The nonparametric nature of this estimator implies that its rate of convergence is no longer the standard $n^{1/2}$. To construct a test for stochastic dominance we extend this strong approximation result to the normalized process $(nh^q)^{1/2}\left(D_{n,\gamma}(z)-g_{\gamma}(z)\right)$. The use of nonparametric kernel estimators for estimating $g_{\gamma}(z)$ renders this extension more difficult to establish; standard results to show the tightness of empirical processes cannot be immediately applied in this context. Instead, we adapt the asymptotic theory developed in Chernozhukov, Lee and Rosen (2012) and Ponomareva (2010), based on strong approximations of nonparametric kernel estimators by a sequence of Brownian bridge processes in an *iid* setting, to a setting with weakly dependent observations. This is done using results by Neumann (1998) on strong approximations of density estimators from weakly dependent observations by density estimators from independent observations.

Note that the sequence containing the information set $I_1,\dots,I_n$ is a weakly dependent time series with a stationary density f^{I_1}. As a counterpart, we consider *iid* random vectors $\widetilde{I}_1,\dots,\widetilde{I}_n$ with the same density f^{I_1} to derive the relevant strong approximation result, see proof of Theorem 1 in appendix and Neumann (1998, pp. 2016 – 2021) for more details on this construction.

Theorem 1. *Let* $\ell_z(\widetilde{I}_t,u_t)=\frac{u_t(y)}{h^{q/2}f^{I_1}(x)}W\left(\frac{\widetilde{I}_t-x}{h}\right)$ *with* $u_t(y)$ *an iid error term obtained from* (8). *Under A.1-A.7,*

$$\sup_{z\in\widetilde{\Omega}}\left|(nh^q)^{1/2}\left(D_{n,\gamma}(z)-g_{\gamma}(z)\right)-G_n(\ell_z)\right| = o_P\left(\delta_n\right)$$

with $z=(x,y)\in\widetilde{\Omega}$, $G_n(\ell_z)$ *a sequence of centered Brownian bridge processes such that* $z\mapsto G_n(\ell_z)$ *has continuous sample paths over* $\widetilde{\Omega}$ *and* δ_n *is such that* $n^{-1/(2q+2)}(h^{-1}log\ n)^{1/2}+(nh^q)^{-1/2}log^{3/2}\ n=o(\delta_n)$.

Let $T_{n,\gamma}=\sup_{z\in\widetilde{\Omega}}(nh^q)^{1/2}D_{n,\gamma}(z)$ be a family of test statistics suitable for testing (5). Under $\widetilde{H}_{0,\gamma}$, Theorem 1 shows that the critical values of the test can be uniformly approximated by the relevant quantiles of the distribution of the supremum of $G_n(\ell_z)$ for n sufficiently large. Let $c_{n,\alpha}$ with $0<\alpha<1$, denote the sequence of $\alpha-$quantiles corresponding to the sequence of approximating distributions.

Theorem 2. *Given Assumptions A.1-A.7,*

(i) *if* $H_{0,\gamma}$ *is true*

$$\lim_{n\to\infty} P\left\{T_{n,\gamma} > c_{n,\alpha}\right\} \le \alpha, \tag{9}$$

with equality under $\widetilde{H}_{0,\gamma}$.

(ii) *if* $H_{0,\gamma}$ *is false*

$$\lim_{n\to\infty} P\left\{T_{n,\gamma} > c_{n,\alpha}\right\} = 1. \tag{10}$$

This theorem shows the consistency of the family of stochastic dominance tests defined by $T_{n,\gamma}$. As a byproduct, condition (i) reveals that for null hypotheses more general than $\widetilde{H}_{0,\gamma}$ the correct asymptotic critical value of the test is smaller than $c_{n,\alpha}$, given n. In this case the test $T_{n,\gamma}$ is undersized producing in turn a loss in statistical power. This problem is still unresolved in the literature; Delgado and Escanciano (2013) partially solve it by proposing conditional stochastic dominance tests in an *iid* setting that are consistent over the boundary of the null hypothesis.

In practice, however, the critical values $c_{n,\alpha}$ are not known and cannot be universally tabulated. The approximation of the distribution of $T_{n,\gamma}$ given by the above family of Brownian bridge processes depends on nuisance parameters as, for example, the density $f^{I_1}(x)$ or the bandwidth parameters h_s if obtained from data-driven methods. Critical values need to be approximated by resampling or simulation methods. The next subsection discusses a simulation method to approximate the p-value of the tests.

3.2 Approximation of the Asymptotic P-Values

The asymptotic distribution of $T_{n,\gamma}$ is nonstandard due to the presence of nuisance parameters. This implies that critical values for stochastic dominance tests of an arbitrary order γ cannot be universally tabulated. In this case there are several alternatives explored in the literature for testing stochastic dominance, namely, simulation and *iid* bootstrap methods as in Barrett and Donald (2003), subsampling and bootstrap as in LMW, and block bootstrap for time series as in Scaillet and Topaloglou (2010). We propose instead a simulation method based on the above nonparametric kernel regression and similar in spirit to the multiplier method for kernels in Chernozhukov, Lee and Rosen (2012)[3].

[3] In a similar nonparametric context, Ponomareva (2010) develops different bootstrap methods for inference on the parameters of unconditional moment inequalities.

We operate conditionally on a realization of the process $\{(Y_t^A, Y_t^B, I_t)\}_{t=1}^n$ denoted by $\omega_n = \{(y_t^A, y_t^B, i_t)\}_{t=1}^n$. Define the simulated process $S_{n,\gamma}^*(z) = (nh^q)^{1/2} D_{n,\gamma}^*(z)$. This process can be generated from

$$D_{n,\gamma}^*(z) = \frac{n^{-1}\sum_{t=1}^{n} d_{t,\gamma}^*(y) W_h\left(\frac{I_t - x}{h}\right)}{\widehat{f}^{I_1}(x)}, \tag{11}$$

with $d_{t,\gamma}^*(y) = d_{t,\gamma}(y)e_t$ and e_t as an external $iid(0,1)$ random variable independent of the data. Interestingly, this process can be interpreted as the Nadaraya-Watson estimator of $g_\gamma^*(I_t, y)$ in the dynamic nonparametric regression

$$d_{t,\gamma}^*(y) = g_\gamma^*(I_t, y) + u_t^*(y), \tag{12}$$

with $u_t^*(y)$ the error term of the nonparametric regression.

Theorem 3. *Under A.1-A.7, the process $S_{n,\gamma}^*(z)$ satisfies that*

$$P_{\omega_n}\left\{\sup_{z\in\widetilde{\Omega}}\left|S_{n,\gamma}^*(z) - \bar{G}_n(\ell_z)\right| > o(\zeta_n)\right\} \to 0 \text{ as } n \to \infty,$$

with $\bar{G}_n(\ell_z)$ an independent and identically distributed copy of the Brownian bridge process $G_n(\ell_z)$ and $\zeta_n \to 0$ as $n \to \infty$; P_{ω_n} denotes the simulated probability conditional on the sample ω_n.

Let $T_{n,\gamma}^* = \sup_{z\in\widetilde{\Omega}} S_{n,\gamma}^*(z)$; this theorem shows the uniform consistency of the simulated critical value obtained from the quantile of the distribution of $T_{n,\gamma}^*$. More formally, under $H_{0,\gamma}$,

$$\lim_{n\to\infty} P_{\omega_n}\left\{T_{n,\gamma}^* > T_{n,\gamma}\right\} \leq \alpha.$$

The distribution of $T_{n,\gamma}^*$ is not directly observed but by operating conditionally on ω_n, it can be approximated to any degree of accuracy. The algorithm to compute the p-value of the test is described as follows.

Algorithm:

1. Construct a grid of $\jmath_1 \times \jmath_2$ points indexed by $z_{ij} = (x_i, y_j)$, with $i = 1, \ldots, \jmath_1$ and $j = 1, \ldots, \jmath_2$ contained in the compact space $A_{\jmath_1\times\jmath_2} \subset \widetilde{\Omega}$; and execute the following steps for $b = 1, \ldots, B$.

2. Generate $\{e_t\}_{t=1}^n$ $iid(0,1)$ random variables independent of the data, and construct $d_{t,\gamma}^*(y_j) = d_{t,\gamma}(y_j)e_t$.

3. Set $D_{n,\gamma}^{*(b)}(z_{ij}) = \frac{n^{-1}\sum_{t=1}^{n} d_{t,\gamma}^*(y_j) W_h\left(\frac{I_t - x_i}{h}\right)}{\widehat{f}^{I_1}(x_i)}$ for all $z_{ij} \in A_{\jmath_1\times\jmath_2}$, with $W_h\left(\frac{I_t - x_i}{h}\right) = \prod_{s=1}^{q} h_s^{-1} w\left(\frac{I_{t,s} - x_{i,s}}{h_s}\right)$;

$w(\cdot)$ is a univariate kernel function for each component of I_t and $h_1, \ldots, h_q$ obtained from optimal cross-validation methods.

4. Set $S^{*(b)}_{n,\gamma}(z_{ij}) = (nh^q)^{1/2}\ D^{*(b)}_{n,\gamma}(z_{ij})$.

5. Store $T^{*(b)}_{n,\gamma} = \sup\limits_{z_{ij} \in A_{J1 \times J2}} S^{*(b)}_{n,\gamma}(z)$.

This algorithm yields a random sample of B observations from the distribution of $\sup\limits_{z \in \widetilde{\Omega}} S^*_{n,\gamma}(z)$. Using the Glivenko-Cantelli theorem and previous assumptions, the empirical p-value conditional on ω_n defined by

$$\widehat{p}^*_{n,B,\gamma} = \frac{1}{B}\sum_{b=1}^{B} 1(T^{*(b)}_{n,\gamma} > T_{n,\gamma})$$

converges in probability to $P_{\omega_n}\left\{T^*_{n,\gamma} > T_{n,\gamma}\right\}$ as $B \to \infty$.

4 Monte-Carlo Simulation Exercise

In this section, we consider two different Monte Carlo simulation experiments to assess the accuracy in finite samples of the nonparametric tests on the first, second and third orders of dynamic conditional stochastic dominance. The first simulation experiment studies the performance of the tests in a cross-sectional regression model, and the second studies the performance of the tests in a simple time series context. For completeness, we also study the power of the test under fixed alternatives.

For the first experiment, the data generating process is

$$Y_i^j = \alpha_0^j + \beta^j X_i + \varepsilon_i^j, \text{ with } j = A, B, \tag{13}$$

with X_i as a univariate $N(0,1)$ random variable. The error term $(\varepsilon^A, \varepsilon^B)$ is a bivariate random variable that follows a standardized Student's-t distribution with $\nu = 30, 5$ degrees of freedom and cross-correlation parameters $\rho(\varepsilon^A, \varepsilon^B) = 0, 0.8$. This distribution is selected to add flexibility to the model and better approximate the behavior of innovations encountered in the modeling of financial time series (see Bollerslev (1987)). The critical values of the different tests are obtained assuming the least favorable case $\widetilde{H}_{0,\gamma}$ under the null hypothesis.

Table 1 reports the empirical size of the stochastic dominance tests $H_{0,\gamma}$ for $\gamma = 1, 2, 3$, for the data generating processes determined by $\alpha_0^A = \alpha_0^B = 0$ and $\beta^A = \beta^B = 0.5$. The results are robust to the choice of degrees of freedom and of the correlation parameter governing the cross-dependence between Y^A and Y^B. The performance of the test improves with the sample size.

Table 2 reports the results on the power of the tests for $H_{0,\gamma}$ for $\gamma = 1, 2, 3$ for two different types of alternative hypotheses. First, we consider the alternative hypothesis given by the stochastic dominance of Y^B over Y^A, characterized by the model parameters $\alpha_0^B = \alpha_0^A + c$ with $c = 0.1, 0.25$, and a bivariate Student's-t distributed error term vector with 30 degrees of freedom and uncorrelated components[4]. The power of the test is slightly higher for the second order compared to the other orders of stochastic dominance under study. It increases with the sample size and as the alternative hypothesis departs from the null hypothesis. The second power analysis in this cross-sectional setting (see Table 2b) assesses whether the test is capable of detecting stochastic efficiency (no dominance of either portfolio). This hypothesis is in the alternative hypothesis to $H_{0,\gamma}$. The following simulation experiment focuses on stochastic efficiency of the first order and is characterized by the processes $Y_i^A = X_i + 0.5\varepsilon_i^A + 0.5\varepsilon_i^B$ and $Y_i^B = X_i + \varepsilon_i^B$, with $X \sim N(0, 1)$, $\rho(\varepsilon^A, \varepsilon^B) = 0$ and $\nu = 30, 5$. The conditional distributions of these variables are both Student's-t distributions with expected values given by the values taken by X. The conditional variance of Y^A is, however, smaller than that of Y^B, implying that for the second and higher orders, these processes are under the null hypothesis $H_{0,\gamma}$. The power of the test for the first order is very high and increases with the sample size. For second and third orders, the empirical rejection probabilities are very close to the nominal sizes.

[INSERT TABLE 1 AND 2 ABOUT HERE]

The simulation section is completed with the study of the size and power of the test for stationary time series processes. The data generating process is

$$Y_t^j = \alpha_0^j + \beta^j (Y_{t-1}^A + Y_{t-1}^B) + \varepsilon_t^j, \text{ with } j = A, B. \tag{14}$$

For $\alpha_0^A = \alpha_0^B = 0$ and $\beta^A = \beta^B = 0.25$, the processes are under the null hypothesis $\widetilde{H}_{0,\gamma}$. The results in the upper panel of Table 3 show that the simulated size of the test is very close to the nominal size for $n = 500$. In contrast to the cross-sectional study, we now observe that for $n = 50$, the simulated size underestimates the true size of the test. This result is more important for first-order stochastic dominance than for second- or third-order stochastic dominance. To assess the size of the test for the general hypothesis $H_{0,\gamma}$ that considers the strict inequality between $\Psi_{I_t,\gamma}^A$ and $\Psi_{I_t,\gamma}^B$, we contemplate two simulation exercises.

[4]The results for other choices of error distribution do not vary qualitatively and are not reported for the sake of of space but are available from the authors upon request.

First, we analyze the size for $\Psi^A_{x,\gamma}(y) - \Psi^B_{x,\gamma}(y) = K$ for all $(x, y) \in \widetilde{\Omega}$ with $K < 0$ constant, and, second, for $\Psi^A_{x,\gamma}(y) - \Psi^B_{x,\gamma}(y) = Ko(n^{-1/2})$. The first experiment is achieved by imposing the model parameters $\alpha^A_0 = 0.1$ and $\alpha^B_0 = 0$, and the second experiment is achieved by using a *local* null hypothesis given by $\alpha^A_0 = 0.1\, n^{-1/2}$ and $\alpha^B_0 = 0$. These results are reported in the middle and lower panels of Table 3. These simulations are consistent with the theory developed above. For the first case, the approximation provided by our simulation method yields an undersized test. For null hypotheses that converge to the least favorable case as n increases, the results improve, and the test yields reasonable estimates of the size for $n = 500$.

Finally, we study the power of the test for alternatives defined by $\alpha^B_0 = \alpha^A_0 + c$ with $c = 0.1, 0.25$. The results are similar to those obtained for the cross-sectional study. The test is consistent under fixed alternatives revealing a nontrivial power in finite samples. Table 4 reports these results[5] for $\alpha^A_0 = 0$.

[INSERT TABLE 3 AND 4 ABOUT HERE]

The good performance of this family of tests in terms of size and power reinforces their usefulness in finite-sample applications.

5 Application: Investment Performance of Sectoral Portfolios

In this section, we apply our nonparametric tests of stochastic dominance to US sectoral portfolios conditional on the dynamics of the market portfolio. The data set consists of monthly excess returns on the ten equally-weighted industry portfolios obtained from the data library on Kenneth French's website and of monthly excess returns on the market portfolio constructed as a value-weighted return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate. The period under study is January 1960 to December 2009. The sectors are Nondurables, Durables, Manufactures, Energy, High Technology, Telecommunications, Shops, Health, Utilities and Others.

Table 5 reports the p-values of the tests of the first, second, and third orders of conditional stochastic dominance. The p-values are obtained assuming the null hypothesis $\widetilde{H}_{0,\gamma}$. Each row in Table 5 shows a vector of simulated p-values that each correspond to the test $H_{0,\gamma} : \Psi^A_{I_t,\gamma} \leq \Psi^B_{I_t,\gamma}$ with portfolio A in the row and portfolio B in the column. If the p-value is higher than the significance level α we cannot reject the null hypothesis. To see if the hypothesis of conditional stochastic dominance of A over B can be accepted, we

[5]To save space the power analysis corresponding to the efficiency between the portfolios is omitted in this time series framework.

need to analyze the p-value of the reverse test given by the dominance of B over A. If the p-value of the latter reverse test is less than the significance level α, we conclude that A dominates B for order γ; otherwise, we cannot differentiate statistically between $\Psi^{A}_{I_t,\gamma}$ and $\Psi^{B}_{I_t,\gamma}$.

The high p-values observed in Table 5 indicate that the null hypothesis of first-stochastic dominance cannot be rejected in either direction for any of the sectors. This result suggests a very similar *distributional* performance of the different sectoral investment portfolios. The study of second order stochastic dominance is more revealing. In this case, we observe that the test between A =High Technology and B =Telecommunications rejects the null hypothesis ($p-$value=0.082), and the reverse test[6] is not rejected ($p-$value=0.220). This means that the portfolio composed of companies working in the Telecommunications sector (Telephone and Television Transmission) has a second-order stochastic dominance over the portfolio of companies in the High-Tech sector (Business Equipment – Computers, Software and Electronic Equipment) conditional on the dynamics of the market portfolio, and, hence, is the choice of risk-averse investors. To check the robustness of this result, we also compute the test for third-order stochastic dominance. The p-value of $\widetilde{H}_{0,3}$ for A =High Technology and B =Telecommunications is 0.098 and 0.796 for the reverse test. The Telecommunications portfolio is also preferred by investors with increasing levels of risk aversion. Similar findings are obtained for the pair A =Shops and B =Telecommunications; the null hypothesis is rejected for the second order ($p-$values=0.072), but the reverse test is not ($p-$value=0.246). For the third-order stochastic dominance, we observe similar results, a p-value of 0.060 for $H_{0,3}$ and a p-value of 0.716 for the corresponding reverse test. This implies that Telecommunications also stochastically dominates the portfolio of companies in the Wholesale, Retail, and Some Services (Laundries, Repair Shops) conditional on the dynamics of the market portfolio.

The overall good performance of Telecommunications compared to the rest of sectors is worth noting. The p-values of the second and third order stochastic dominance tests of this portfolio against the rest of the sectoral portfolios (the column corresponding to Telecommunications in the middle and lower panels) take values of approximately 0.12-0.15. At the same time, the p-values in the row corresponding to this sector are all higher than 0.20. These combined results show that at the 20% significance level, Durables, Nondurables, Manufactures, High-Technology and Shops are dominated by Telecommunications for the second and third orders, conditional on the dynamics of the market portfolio.

To assess the importance of considering the dynamics of the market portfolio we repeat the empirical

[6] The alternative hypothesis in this case is the dominance of A by B. The test statistic is constructed by reversing the roles of A and B. Critical values are obtained assuming $\widetilde{H}_{0,\gamma}$. The p-values are calculated from the algorithm in Section 3.2.

exercise by implementing unconditional versions of the various stochastic dominance tests. The very weak serial correlation between the monthly excess returns on the sectoral portfolios implies that we can apply the method developed by Barrett and Donald (2003) through accommodating cross-dependence between portfolios A and B to approximate the p-values of the different stochastic dominance tests. Table 6 reports the p-values of the different tests; the results suggest a completely different ordering of investment strategies. The Durables sector is, for example, an efficient portfolio in the sense that is not dominated nor dominates any other sectoral portfolio; similarly, High Technology is stochastically efficient for the first order but not for higher orders. In fact, as for the conditional case, we observe that Telecommunications dominates High Technology for orders higher than one. Finally, it is worth noting that in this unconditional setting Utilities performs very well compared to the rest of the sectoral portfolios. The p-values of the test for the second order reveal that Utilities dominates every other sector except for Nondurables and Telecommunications. The latter two sectors dominate Utilities.

6 Concluding Remarks

For the concept of stochastic dominance to be fully operational, it has to be exploited dynamically. While there are many influential methods for testing the hypothesis of stochastic dominance in an unconditional or marginal setting, there are only a few methods that aim to do this dynamically or conditionally on an information set. Moreover, these conditional stochastic dominance tests rely heavily on assuming an appropriate parametric structure for the dependence between the variables and hence are subject to misspecification issues.

This paper presents a nonparametric test for conditional stochastic dominance that easily accommodates the presence of dynamics in the variables without having to impose strong assumptions on the specific form of these dynamics. The asymptotic theory of the test is simple, and p-values can be approximated by simulation methods. The test has good finite-sample performance and is easy to implement under a variety of conditional settings. The application to studying investment performance on sectoral indices shows that the Telecommunications sector dominates the High-Tech and Shop sectors for the second and third orders of stochastic dominance. Furthermore, at the 20% significance level, this portfolio also dominates for second and third orders of stochastic dominance the other sectoral portfolios. The advantage of the Telecommunications sector compared to the rest of portfolios appears to be in the low volatility of its returns for a given expected return level. These results are also observed for increasing risk aversion.

Universidad Carlos III de Madrid, Spain; University of Southampton, U.K.

APPENDIX

Proof of Theorem 1: To prove this theorem we follow Theorem 8 in Chernozhukov, Lee and Rosen (2012). The theorem developed by these authors is derived in an *iid* setting so we need to transform our stationary weakly dependent framework into *iid*. This is done by applying the results in Neumann (1998) to the stationary Bahadur representation of $(nh^q)^{1/2}(D_{n,\gamma}(z) - g_\gamma(z))$ in terms of $u_t(y)$ and $W(\frac{I_t-x}{h})$. First, we derive this asymptotic expansion for the stationary case. Note that

$$(nh^q)^{1/2}(D_{n,\gamma}(z) - g_\gamma(z)) =$$

$$\frac{1}{(nh^q)^{1/2}\widehat{f}^{I_1}(x)}\left(\sum_{t=1}^{n}(g_\gamma(I_t,y) - g_\gamma(z))W\left(\frac{I_t-x}{h}\right)\right) + \frac{1}{(nh^q)^{1/2}\widehat{f}^{I_1}(x)}\sum_{t=1}^{n}u_t(y)W\left(\frac{I_t-x}{h}\right)$$

with $W(\cdot) = h^q W_h(\cdot)$ and $z = (x,y)$. The expression $\frac{1}{(nh^q)^{1/2}\widehat{f}^{I_1}(x)}\left(\sum_{t=1}^{n}(g_\gamma(I_t,y) - g_\gamma(z))W\left(\frac{I_t-x}{h}\right)\right)$ converges to zero uniformly over $z \in \widetilde{\Omega}$. This is a consequence of the uniform convergence of nonparametric kernel estimators derived in Masry (1996) under strong mixing conditions. These results can be applied in our setting by imposing a beta mixing condition (A.1) limiting the extent of serial dependence in the data, the smoothness of the function $g_\gamma(z)$ and the Lipschitz conditions in A.3 and A.4. Then

$$(nh^q)^{1/2}(D_{n,\gamma}(z) - g_\gamma(z)) = \frac{1}{(nh^q)^{1/2}f^{I_1}(x)}\sum_{t=1}^{n}u_t(y)W\left(\frac{I_t-x}{h}\right) + o_P(1),$$

uniformly over $\widetilde{\Omega}$.

To obtain the characterization of this Bahadur representation in terms of an *iid* process we use the following;

$$(nh^q)^{1/2}(D_{n,\gamma}(z) - g_\gamma(z)) =$$

$$\frac{1}{(nh^q)^{1/2}f^{I_1}(x)}\sum_{t=1}^{n}u_t(y)\left(W\left(\frac{I_t-x}{h}\right) - W\left(\frac{\widetilde{I}_t-x}{h}\right)\right) + \frac{1}{(nh^q)^{1/2}f^{I_1}(x)}\sum_{t=1}^{n}u_t(y)W\left(\frac{\widetilde{I}_t-x}{h}\right) + o_P(1)$$

In order for

$$\sup_{z\in\widetilde{\Omega}}\left|(nh^q)^{1/2}(D_{n,\gamma}(z) - g_\gamma(z)) - \frac{1}{(nh^q)^{1/2}f^{I_1}(x)}\sum_{t=1}^{n}u_t(y)W\left(\frac{\widetilde{I}_t-x}{h}\right)\right| \tag{15}$$

to converge to zero as $n \to \infty$ it suffices that

$$\sup_{z \in \widetilde{\Omega}} \left| \frac{1}{(nh^q)^{1/2} f^{I_1}(x)} \sum_{t=1}^{n} u_t(y) \left(W\left(\frac{I_t - x}{h}\right) - W\left(\frac{\widetilde{I}_t - x}{h}\right)\right) \right| \tag{16}$$

to be $o_P(1)$ as $n \to \infty$.

Before proceeding with the proof of this condition we briefly discuss how to construct the *iid* sequence $\widetilde{I}_t$. The method developed by Neumann (1998) establishes a link between density estimation under weak dependence and density estimation based on independent observations by embedding the random variables, I_t and $\widetilde{I}_t$ in our case, in a common marked Poisson process N indexed by time as well as spatial position. More specifically, the Poisson process N is defined on $(0, \infty) \times \Omega'$ with intensity function equal to the density function $f^{I_1}(x)$ for $x \in \Omega'$. Neumann (1998, pp 2018 – 2021) describes the method to generate copies of the observations $\{I_1, \ldots, I_n\}$ and $\{\widetilde{I}_1, \ldots, \widetilde{I}_n\}$ retaining the joint distribution of the original random vector $\{I_1, \ldots, I_n\}$. Since the transition densities are usually different from the stationary density, the construction for the time series model borrows some probability mass assigned to future time points in the *iid* model. This is the reason to introduce a time axis for the Poisson process embedding method.

The proof of Theorem 1 does not require specific knowledge on the construction of the *iid* process. To complete the proof we must prove condition (16). Note that the process $d_{t,\gamma}(y)$ is bounded over a compact set implying in turn that the error process $u_t(y)$ is also bounded. Let $C_4,\ C_5$ with $-\infty < C_4 < 0 < C_5 < \infty$ be universal lower and upper bounds, respectively, of the process $u_t(y)$. Then, for each $x \in \Omega'$ the expression inside the summation operator can be reordered in terms of positive and negative values as

$$\sum_{t=1}^{k(x)} u_t(y) \left(W\left(\frac{I_t - x}{h}\right) - W\left(\frac{\widetilde{I}_t - x}{h}\right)\right) + \sum_{t=k(x)+1}^{n} u_t(y) \left(W\left(\frac{I_t - x}{h}\right) - W\left(\frac{\widetilde{I}_t - x}{h}\right)\right)$$

with $k(x)$ denoting the number of terms with $W\left(\frac{I_t - x}{h}\right) - W\left(\frac{\widetilde{I}_t - x}{h}\right) < 0$ and $n - k(x)$ the number of remaining observations; note that by construction $\sup_{x \in \Omega'} k(x) = O(n)$. It is trivial to see that, for each $x \in \Omega'$, this expression can be upper bounded by

$$\sum_{t=1}^{k(x)} C_4 \left(W\left(\frac{I_t - x}{h}\right) - W\left(\frac{\widetilde{I}_t - x}{h}\right)\right) + \sum_{t=k(x)+1}^{n} C_5 \left(W\left(\frac{I_t - x}{h}\right) - W\left(\frac{\widetilde{I}_t - x}{h}\right)\right).$$

After suitable algebra mainly consisting of adding and substracting from the preceding expression the term $\sum_{t=k(x)+1}^{n} C_4 \left(W\left(\frac{I_t - x}{h}\right) - W\left(\frac{\widetilde{I}_t - x}{h}\right)\right)$, for each $x \in \Omega'$, the expression inside the supremum functional in (16) is upper bounded by

$$\frac{|C_4|}{f^{I_1}(x)}(nh^q)^{1/2}\big|\widehat{f}_n^{I_1}(x)-\widehat{f}_n^{\widetilde{I}_1}(x)\big|+\frac{|C_5-C_4|}{f^{I_1}(x)}\frac{(n-k(x))h^{q/2}}{n^{1/2}}\big|\widehat{f}_{n-k(x)}^{I_1}(x)-\widehat{f}_{n-k(x)}^{\widetilde{I}_1}(x)\big| \tag{17}$$

where $\widehat{f}_n^{\widetilde{I}_1}(x)=\frac{1}{nh^q}\sum_{t=1}^{n}W\left(\frac{\widetilde{I}_t-x}{h}\right)$ is the kernel estimator of $f^{I_1}(x)$ corresponding to the associated *iid* process $\widetilde{I}_t$ and $\widehat{f}_n^{I_1}(x)$ the stationary kernel estimator counterpart. The subscript n in $\widehat{f}$ refers to the number of observations involved in the estimation of the density functions.

Taking the supremum of (17) over $x \in \Omega'$ and using that $f^{I_1}(x)$ is bounded away from zero, we can upper bound the previous expression as

$$C_6(nh^q)^{1/2}\sup_{x\in\Omega'}\big|\widehat{f}_n^{I_1}(x)-\widehat{f}_n^{\widetilde{I}_1}(x)\big|+C_7(nh^q)^{1/2}\sup_{x_o\in\Omega'}\left(\sup_{x\in\Omega'}\big|\widehat{f}_{n-k(x_o)}^{I_1}(x)-\widehat{f}_{n-k(x_o)}^{\widetilde{I}_1}(x)\big|\right) \tag{18}$$

where C_6 and C_7 are suitable positive constants. Neumann (1998) shows under some regularity conditions, mainly A.1 and A.2, that

$$\sup_{x\in\Omega'}\big|\widehat{f}_n^{I_1}(x)-\widehat{f}_n^{\widetilde{I}_1}(x)\big|=O\left(n^{-1/2}log\ n\right).$$

Then, expression (18) reads as

$$C_6\ O\left(h^{q/2}log\ n\right)+C_7\sup_{x_o\in\Omega'}O\left(\left(\frac{n}{n-k(x_o)}\right)^{1/2}h^{q/2}log\ (n-k(x_o))\right). \tag{19}$$

Then, it is not difficult to see that under $A.5$, more specifically under condition $h^{q/2}log\ n \to 0$, and using that $n/(n-k(x_o))$ converges to a constant as $n \to \infty$, expression (19) converges to zero as $n \to \infty$.

To complete the proof of the theorem we apply Theorem 8 in Chernozhukov, Lee and Rosen (2012) to

$$\frac{1}{(nh^q)^{1/2}f^{I_1}(x)}\sum_{t=1}^{n}u_t(y)W\left(\frac{\widetilde{I}_t-x}{h}\right)+o_P(1).$$

As a result, it holds that

$$\sup_{z\in\widetilde{\Omega}}\big|(nh^q)^{1/2}\left(D_{n,\gamma}(z)-g_\gamma(z)\right)-G_n(\ell_z)\big|=o_P\left(\delta_n\right)$$

with $G_n(\ell_z)$ a Brownian bridge process, $\ell_z(\widetilde{I}_t,u_t)=\frac{u_t(y)}{h^{q/2}f^{I_1}(x)}W\left(\frac{\widetilde{I}_t-x}{h}\right)$ and δ_n a sequence satisfying that $n^{-1/(2q+2)}(h^{-1}log\ n)^{1/2}+(nh^q)^{-1/2}log^{3/2}\ n=o(\delta_n)$.

Proof of Theorem 2: Let $c_{n,\alpha}$ be the sequence of critical values at a significance level α obtained from the distribution of the supremum of $G_n(\ell_z)$. Under A.1-A.7, the strong approximation in Theorem 1 implies that under $H_{0,\gamma}$,

$$\lim_{n\to\infty} P\left\{T_{n,\gamma} > c_{n,\alpha}\right\} \leq \alpha.$$

Under $\widetilde{H}_{0,\gamma}$ the distribution of $(nh^q)^{1/2} D_{n,\gamma}(z)$ is uniformly approximated by the above sequence of Brownian bridge processes for n sufficiently large. Then, it follows that the critical value of $T_{n,\gamma}$ is uniformly consistently approximated by $c_{n,\alpha}$ for n sufficiently large, giving the equality in (9).

To prove Condition (ii), note that Theorem 1 shows that the process $(nh^q)^{1/2}(D_{n,\gamma}(z) - g_\gamma(z))$ is uniformly approximated by the above sequence of Brownian bridge processes. The distribution of $(nh^q)^{1/2} \sup_{z\in\widetilde{\Omega}}(D_{n,\gamma}(z) - g_\gamma(z))$ converges to the distribution of the supremum of $G_n(\ell_z)$ uniformly over $z \in \widetilde{\Omega}$. If $H_{0,\gamma}$ is false, this process is majorized in distribution by the process $(nh^q)^{1/2} \sup_{z\in\widetilde{\Omega}} D_{n,\gamma}(z)$ that diverges to infinity since $nh^q \to \infty$ as $n \to \infty$. Hence, it is immediate to see that

$$\lim_{n\to\infty} P\left\{T_{n,\gamma} > c_{n,\alpha}\right\} = 1.$$

Proof of Theorem 3: By construction,

$$S^*_{n,\gamma}(z) = \frac{1}{(nh^q)^{1/2} \widehat{f}^{I_1}(x)} \sum_{t=1}^{n} d^*_{t,\gamma}(y) W\left(\frac{I_t - x}{h}\right). \tag{20}$$

Using the same steps as for the proof of Theorem 1, it is not difficult to see that under the null hypothesis $\widetilde{H}_{0,\gamma}$ and assumptions A.1-A.7 the process $S^*_{n,\gamma}(z)$ has the following Bahadur representation:

$$S^*_{n,\gamma}(z) = \frac{1}{(nh^q)^{1/2} f^{I_1}(x)} \sum_{t=1}^{n} u^*_t(y) W\left(\frac{\widetilde{I}_t - x}{h}\right) + o_P(1),$$

with $\widetilde{I}_t$ the counterpart *iid* random vector of the weakly dependent sequence I_t and $u^*_t(y) = u_t(y)e_t$ with e_t an $iid(0,1)$ random variable independent of the data.

The simulated process $S^*_{n,\gamma}(z)$ can be expressed as $\frac{1}{n^{1/2}}\sum_{t=1}^{n} e_t \ell_z(\widetilde{I}_t, u_t)$, and Theorem 9 in Chernozhukov, Lee and Rosen (2012) can be applied to obtain the result in Theorem 3.

References

Anderson, G.J., "Nonparametric Tests of Stochastic Dominance in Income Distributions," Econometrica 64

(1996), 1183-1193.

Barrett, G. and S. Donald, “Consistent Tests for Stochastic Dominance,” Econometrica 71 (2003), 71-104.

Bollerslev, T, “A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return,” The Review of Economics and Statistics 69 (1987), 542-547.

Buja, A., T. Hastie, and R. Tibshirani, “Linear Smoothers and Additive Models,” Annals of Statistics 17 (1989), 453-555.

Chernozhukov, V., S. Lee, and A. Rosen, “Intersection Bounds: Estimation and Inference,” Forthcoming in Econometrica.

Davidson, R., and J.Y. Duclos, “Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality,” Econometrica 68 (2000), 1435-1464.

Delgado, M.A., and J.C. Escanciano, “Nonparametric Tests for Conditional Symmetry in Dynamic Models,” Journal of Econometrics 141 (2007), 652-682.

Delgado, M.A., and J.C. Escanciano, “Conditional Stochastic Dominance Testing,” Journal of Business and Economic Statistics 31 (2013), 16-28.

Fishburn, P.C., “Mean-Risk Analysis with Risk Associated with Below-Target Returns,” The American Economic Review 67 (1977), 116-126.

Hansen, B.E., “Inference when a Nuisance Parameter is not Identified Under the Null Hypothesis,” Econometrica 64 (1996), 413-430.

Hansen, B.E., “Uniform Convergence Rates for Kernel Estimation with Dependent Data,” Econometric Theory 24 (2008), 726-748.

Hastie, T., and R. Tibshirani, Generalized Additive Models, London: Chapman and Hall, 1990.

Kaur, A., B. L. Prakasa Rao, and H. Singh, “Testing for Second-Order Stochastic Dominance of Two Distributions,” Econometric Theory 10 (1994), 849-866.

Klecan, L., R. McFadden and R. McFadden, “A Robust Test for Stochastic Dominance,” Working Paper, Department of Economics, MIT, 1991.

Kopa, M., and T. Post, "A Portfolio Optimality Test based on the First-Order Stochastic Dominance Criterion," Journal of Financial and Quantitative Analysis 44 (2009), 1103-1124.

Larsen, G.A., and B.G. Resnick, "Bootstrapping a Distance Test for Stochastic Dominance Analysis," Review of Quantitative Finance and Accounting 3 (1993), 61-69.

Levy, H., Stochastic Dominance: Investment Decision Making under Uncertainty. 2^{nd} ed. (New York: Springer, 2006).

Li, Q., and J.S. Racine, Nonparametric Econometrics. Theory and Practice. (New Jersey: Princeton University Press, 2007).

Linton, O., E. Maasoumi, and Y.J. Whang, "Consistent Testing for Stochastic Dominance under General Sampling Schemes," Review of Economic Studies 72 (2005), 735-765.

Linton, O., and J.P. Nielsen, "A kernel method of estimating structured nonparametric regression based on marginal integration," Biometrika 82 (1995), 93-100.

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

Masry, E., "Multivariate Local Polynomial Regression for Time Series: Uniform Strong Consistency and Rates," Journal of Time Series Analysis 17 (1996), 571-599.

McFadden, D., "Testing for Stochastic Dominance," In Studies in the Economics of Uncertainty, edited by T.B. Fomby and T.K. Seo. (New York: Springer-Verlag, 1989).

Nadaraya, E. A., "On nonparametric estimates of density functions and regression curves," Theory of Applied Probability 10 (1965), 186-190.

Neumann, M.H., "Strong approximation of density estimators from weakly dependent observations by density estimators from independent observations," The Annals of Statistics 26 (1998), 2014-2048.

Newey, W.K., "Kernel estimation of partial means and a general variance estimator," Econometric Theory 10 (1994), 233-253.

Ponomareva, M., "Inference in Models Defined by Conditional Moment Inequalities with Continuous Covariates," Working Paper, University of Western Ontario, 2010.

Porter, R.B., "Semivariance and Stochastic Dominance: A Comparison," American Economic Review 64 (1974), 200-204.

Post, T., "Empirical Tests for Stochastic Dominance Efficiency," Journal of Finance 58 (2003), 1905-1932.

Romano, J.P., and M. Wolf, "Alternative Tests for Monotonicity in Expected Asset Returns," Technical Report No. $2011 - 06$. Department of Statistics, Standford University, 2011.

Scaillet, O., and N. Topaloglou, "Testing for Stochastic Dominance Efficiency," Journal of Business and Economic Statistics 28 (2010), 169-180.

Shalit, H., and S. Yitzhaki, "Marginal Conditional Stochastic Dominance," Management Science 40 (1994), 670-684.

Stone, B.K., "A General Class of Three-Parameter Risk Measures," Journal of Finance 28 (1973), 675-685.

Watson, G.S., "Smooth Regression Analysis," Sankhya 26:15 (1964), 359-372.

Table 1. Empirical size for $H_{0,\gamma}$, $\gamma = 1, 2, 3$ with $Y_i^j = \alpha_0^j + \beta^j X_i + \varepsilon_i^j$, $i = 1, \ldots, n$, $\varepsilon_i^j \sim t_\nu$, $j = A, B$ and $\nu = 30, 5$. $\alpha_0^A = \alpha_0^B = 0$ and $\beta^A = \beta^B = 0.5$. $m = 500$ Monte Carlo simulations.

	$\gamma = 1$				$\gamma = 2$				$\gamma = 3$			
$\alpha \backslash n$	50	500	2000	5000	50	500	2000	5000	50	500	2000	5000
$\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.066	0.074	0.118	0.078	0.074	0.116	0.116	0.104	0.076	0.098	0.122	0.072
0.05	0.022	0.042	0.060	0.046	0.030	0.060	0.056	0.050	0.042	0.062	0.062	0.044
0.01	0.002	0.008	0.010	0.008	0.006	0.010	0.012	0.010	0.008	0.012	0.018	0.010
$\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0.8$												
0.10	0.016	0.118	0.066	0.078	0.030	0.106	0.102	0.092	0.050	0.142	0.088	0.088
0.05	0.000	0.064	0.038	0.030	0.002	0.056	0.042	0.038	0.002	0.062	0.034	0.040
0.01	0.000	0.012	0.006	0.008	0.000	0.010	0.006	0.010	0.000	0.016	0.006	0.010
$\nu = 5$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.080	0.086	0.116	0.120	0.082	0.092	0.122	0.118	0.088	0.090	0.114	0.116
0.05	0.024	0.038	0.056	0.062	0.030	0.034	0.068	0.048	0.030	0.040	0.060	0.058
0.01	0.002	0.002	0.006	0.012	0.000	0.002	0.018	0.012	0.002	0.004	0.018	0.014
$\nu = 5$ and $\rho(\varepsilon^A, \varepsilon^B) = 0.8$												
0.10	0.030	0.086	0.102	0.090	0.076	0.094	0.106	0.140	0.100	0.084	0.076	0.118
0.05	0.004	0.040	0.048	0.044	0.028	0.044	0.054	0.058	0.030	0.038	0.034	0.076
0.01	0.000	0.006	0.010	0.004	0.002	0.004	0.010	0.008	0.000	0.004	0.008	0.014

Table 2a. Empirical power for $H_{0,\gamma}$, $\gamma = 1, 2, 3$ with $Y_i^j = \alpha_0^j + \beta^j X_i + \varepsilon_i^j$, $i = 1, \dots, n$, $\varepsilon_i^j \sim t_\nu$, $j = A, B$ and $\nu = 30$. $m = 500$ Monte Carlo simulations.

	$\gamma = 1$				$\gamma = 2$				$\gamma = 3$			
$\alpha \backslash n$	50	500	2000	5000	50	500	2000	5000	50	500	2000	5000
$\alpha_0^A = 0$, $\alpha_0^B = 0.1$ and $\beta^A = \beta^B = 0.5$; $\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.126	0.230	0.244	0.508	0.168	0.272	0.356	0.692	0.144	0.250	0.316	0.504
0.05	0.040	0.120	0.132	0.340	0.068	0.150	0.202	0.472	0.068	0.136	0.182	0.348
0.01	0.004	0.032	0.032	0.112	0.006	0.046	0.050	0.158	0.004	0.042	0.054	0.120
$\alpha_0^A = 0$, $\alpha_0^B = 0.25$ and $\beta^A = \beta^B = 0.5$; $\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.186	0.432	0.692	1.000	0.312	0.706	0.958	1.000	0.292	0.540	0.830	0.998
0.05	0.072	0.246	0.476	0.988	0.142	0.470	0.788	1.000	0.130	0.342	0.572	0.988
0.01	0.006	0.086	0.140	0.766	0.024	0.114	0.326	0.990	0.016	0.080	0.220	0.762

Table 2b. Empirical power for $H_{0,\gamma}$, $\gamma = 1, 2, 3$ with $Y_i^A = X_i + 0.5\varepsilon_i^A + 0.5\varepsilon_i^B$ and $Y_i^B = X_i + \varepsilon_i^B$, where $\varepsilon_i^j \sim t_\nu$, $\nu = 30, 5$ and $j = A, B$. $m = 500$ Monte Carlo simulations.

	$\gamma = 1$				$\gamma = 2$				$\gamma = 3$			
$\alpha \backslash n$	50	500	2000	5000	50	500	2000	5000	50	500	2000	5000
$\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.072	0.228	0.428	0.998	0.078	0.080	0.094	0.066	0.072	0.028	0.032	0.006
0.05	0.022	0.120	0.248	0.940	0.026	0.028	0.056	0.040	0.030	0.012	0.012	0.002
0.01	0.000	0.020	0.068	0.540	0.004	0.010	0.016	0.008	0.002	0.002	0.002	0.000
$\nu = 5$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.036	0.186	0.694	0.996	0.040	0.060	0.064	0.078	0.052	0.046	0.016	0.012
0.05	0.006	0.106	0.402	0.944	0.006	0.038	0.034	0.048	0.012	0.022	0.008	0.004
0.01	0.000	0.008	0.110	0.466	0.002	0.010	0.006	0.018	0.000	0.002	0.000	0.000

Table 3. Empirical size for $H_{0,\gamma}$, $\gamma = 1, 2, 3$ with $Y_t^j = \alpha_0^j + \beta^j (Y_{t-1}^A + Y_{t-1}^B) + \varepsilon_t^j$, $t = 1, \ldots, n$, $\varepsilon_t^j \sim t_\nu$, $j = A, B$ and $\nu = 30$. $m = 500$ Monte Carlo simulations.

	$\gamma = 1$				$\gamma = 2$				$\gamma = 3$			
$\alpha \setminus n$	50	500	2000	5000	50	500	2000	5000	50	500	2000	5000
$\alpha_0^A = \alpha_0^B = 0$, and $\beta^A = \beta^B = 0.25$; $\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.086	0.096	0.092	0.096	0.000	0.100	0.086	0.100	0.000	0.102	0.092	0.102
0.05	0.036	0.050	0.040	0.038	0.000	0.052	0.050	0.054	0.000	0.050	0.044	0.054
0.01	0.006	0.016	0.012	0.012	0.000	0.014	0.014	0.008	0.000	0.010	0.008	0.006
$\alpha_0^A = 0.1$, $\alpha_0^B = 0$ and $\beta^A = \beta^B = 0.25$; $\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.040	0.040	0.014	0.006	0.046	0.014	0.016	0.002	0.060	0.024	0.018	0.000
0.05	0.010	0.014	0.002	0.002	0.018	0.010	0.004	0.000	0.026	0.018	0.008	0.000
0.01	0.002	0.002	0.000	0.000	0.000	0.002	0.000	0.000	0.002	0.002	0.000	0.000
$\alpha_0^A = 0.1/n^{1/2}$, $\alpha_0^B = 0$ and $\beta^A = \beta^B = 0.25$; $\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.008	0.044	0.050	0.064	0.034	0.058	0.066	0.048	0.034	0.062	0.072	0.060
0.05	0.002	0.022	0.028	0.026	0.004	0.018	0.036	0.020	0.008	0.024	0.034	0.036
0.01	0.000	0.002	0.010	0.008	0.000	0.004	0.010	0.004	0.000	0.006	0.008	0.006

Table 4. Empirical power for $H_{0,\gamma}$, $\gamma = 1, 2, 3$ with $Y_t^j = \alpha_0^j + \beta^j (Y_{t-1}^A + Y_{t-1}^B) + \varepsilon_t^j$, $t = 1, \ldots, n$, $\varepsilon_t^j \sim t_\nu$, $j = A, B$ and $\nu = 30$. $m = 500$ Monte Carlo simulations.

	$\gamma = 1$				$\gamma = 2$				$\gamma = 3$			
$\alpha \setminus n$	50	500	2000	5000	50	500	2000	5000	50	500	2000	5000
$\alpha_0^A = 0$, $\alpha_0^B = 0.1$ and $\beta^A = \beta^B = 0.25$; $\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.096	0.268	0.426	0.786	0.140	0.358	0.530	0.930	0.146	0.300	0.456	0.816
0.05	0.032	0.136	0.236	0.622	0.054	0.202	0.384	0.812	0.058	0.164	0.288	0.612
0.01	0.000	0.030	0.072	0.256	0.010	0.048	0.132	0.420	0.006	0.042	0.100	0.266
$\alpha_0^A = 0$, $\alpha_0^B = 0.25$ and $\beta^A = \beta^B = 0.25$; $\nu = 30$ and $\rho(\varepsilon^A, \varepsilon^B) = 0$												
0.10	0.176	0.662	0.928	1.000	0.290	0.834	0.990	1.000	0.274	0.704	0.938	1.000
0.05	0.070	0.436	0.798	1.000	0.124	0.642	0.948	1.000	0.120	0.488	0.804	1.000
0.01	0.004	0.134	0.384	0.972	0.010	0.274	0.638	1.000	0.012	0.180	0.420	0.968

Table 5. Empirical p-values for $H_{0,\gamma} : \Psi^A_{I_t,\gamma} \leq \Psi^B_{I_t,\gamma}$ with $\gamma = 1, 2, 3$. The market portfolio is a value-weighted return all NYSE, AMEX, and NASDAQ stocks minus the one-month Treasury bill rate. The period under study is January 1960 to December 2009.

A \ B	Nondurables	Durables	Manuf	Energy	High Tech.	Telecom	Shops	Health	Utilities	Others
					$\gamma = 1$					
Nondurables	–	0.690	0.366	0.372	0.722	0.512	0.764	0.436	0.380	0.404
Durables	0.380	–	0.294	0.454	0.444	0.410	0.390	0.430	0.408	0.406
Manuf.	0.354	0.528	–	0.406	0.434	0.424	0.320	0.402	0.422	0.390
Energy	0.456	0.534	0.630	–	0.634	0.392	0.456	0.324	0.432	0.476
High Tech.	0.404	0.756	0.372	0.440	–	0.396	0.414	0.404	0.386	0.428
Telecom	0.422	0.478	0.372	0.386	0.444	–	0.454	0.378	0.414	0.434
Shops	0.372	0.662	0.254	0.388	0.410	0.196	–	0.464	0.398	0.418
Health	0.696	0.610	0.576	0.404	0.556	0.572	0.692	–	0.380	0.568
Utilities	0.388	0.620	0.416	0.338	0.412	0.642	0.586	0.276	–	0.384
Others	0.390	0.522	0.534	0.278	0.398	0.422	0.378	0.256	0.404	–
					$\gamma = 2$					
Nondurables	–	0.722	0.408	0.540	0.324	0.170	0.472	0.448	0.142	0.574
Durables	0.176	–	0.278	0.310	0.264	0.148	0.178	0.270	0.232	0.238
Manuf.	0.128	0.832	–	0.540	0.444	0.128	0.328	0.258	0.174	0.394
Energy	0.346	0.642	0.298	–	0.350	0.268	0.586	0.344	0.250	0.388
High Tech.	0.092	0.614	0.108	0.220	–	**0.082**	0.218	0.406	0.160	0.240
Telecom	0.280	0.472	0.302	0.230	0.220	–	0.246	0.238	0.152	0.326
Shops	0.150	0.670	0.236	0.218	0.430	**0.072**	–	0.306	0.146	0.246
Health	0.178	0.554	0.324	0.608	0.250	0.612	0.328	–	0.280	0.320
Utilities	0.596	0.658	0.512	0.488	0.466	0.720	0.502	0.500	–	0.548
Others	0.244	0.774	0.250	0.178	0.514	0.276	0.350	0.322	0.140	–
					$\gamma = 3$					
Nondurables	–	0.874	0.662	0.750	0.764	0.078	0.688	0.264	0.142	0.808
Durables	0.134	–	0.142	0.132	0.158	0.138	0.168	0.194	0.182	0.160
Manuf.	0.128	1.000	–	0.238	0.462	0.128	0.142	0.162	0.240	0.486
Energy	0.354	0.910	0.736	–	0.708	0.128	0.714	0.288	0.244	0.776
High Tech.	0.134	0.474	0.120	0.162	–	**0.098**	0.136	0.216	0.170	0.144
Telecom	0.490	0.828	0.754	0.752	0.796	–	0.716	0.226	0.164	0.864
Shops	0.148	0.912	0.674	0.536	0.858	**0.060**	–	0.256	0.150	0.782
Health	0.278	0.748	0.572	0.732	0.716	0.438	0.550	–	0.160	0.680
Utilities	0.782	0.810	0.722	0.906	0.754	0.674	0.688	0.816	–	0.728
Others	0.226	0.736	0.188	0.140	0.258	0.154	0.186	0.212	0.150	–

Table 6. Empirical p-values for the unconditional test $\Psi_\gamma^A \leq \Psi_\gamma^B$ with $\gamma = 1, 2, 3$. The period under study is January 1960 to December 2009.

A \ B	Nondurables	Durables	Manuf	Energy	High Tech.	Telecom	Shops	Health	Utilities	Others
					$\gamma = 1$					
Nondurables	–	0.014	0.294	0.100	0.000	0.690	0.018	0.264	0.934	0.032
Durables	0.000	–	0.020	0.120	0.272	0.018	0.064	0.038	0.004	0.026
Manuf.	0.044	0.020	–	0.160	0.002	0.836	0.074	0.520	0.218	0.234
Energy	0.020	0.548	0.748	–	0.052	0.440	0.774	0.726	0.100	0.362
High Tech.	0.000	0.222	0.000	0.016	–	0.002	0.006	0.004	0.000	0.006
Telecom	0.134	0.060	0.198	0.048	0.000	–	0.048	0.050	0.420	0.022
Shops	0.004	0.562	0.600	0.650	0.054	0.386	–	0.458	0.148	0.322
Health	0.016	0.308	0.910	0.378	0.006	0.732	0.586	–	0.252	0.700
Utilities	0.030	0.002	0.028	0.000	0.000	0.274	0.004	0.018	–	0.002
Others	0.004	0.248	0.240	0.568	0.006	0.220	0.328	0.568	0.048	–
					$\gamma = 2$					
Nondurables	–	0.970	0.968	0.660	0.980	0.790	0.962	0.886	0.512	0.978
Durables	0.000	–	0.000	0.046	0.282	0.016	0.020	0.012	0.000	0.016
Manuf.	0.002	0.980	–	0.188	0.610	0.436	0.236	0.304	0.046	0.644
Energy	0.052	0.978	0.558	–	0.970	0.392	0.938	0.548	0.044	0.968
High Tech.	0.000	0.140	0.000	0.010	–	0.002	0.004	0.000	0.000	0.002
Telecom	0.056	0.978	0.486	0.114	0.456	–	0.240	0.240	0.112	0.442
Shops	0.002	0.946	0.310	0.434	0.936	0.230	–	0.266	0.040	0.806
Health	0.024	0.964	0.862	0.476	0.976	0.466	0.866	–	0.090	0.978
Utilities	0.076	0.952	0.650	0.132	0.516	0.880	0.276	0.278	–	0.526
Others	0.000	0.978	0.116	0.272	0.740	0.112	0.300	0.214	0.004	–
					$\gamma = 3$					
Nondurables	–	0.924	0.912	0.748	0.964	0.708	0.906	0.784	0.486	0.966
Durables	0.000	–	0.000	0.014	0.648	0.032	0.010	0.012	0.014	0.024
Manuf.	0.006	0.892	–	0.268	0.828	0.426	0.478	0.256	0.184	0.680
Energy	0.170	0.938	0.894	–	0.944	0.438	0.900	0.532	0.194	0.946
High Tech.	0.000	0.250	0.008	0.008	–	0.022	0.014	0.010	0.008	0.026
Telecom	0.044	0.936	0.908	0.290	0.946	–	0.622	0.300	0.202	0.940
Shops	0.006	0.860	0.378	0.362	0.798	0.304	–	0.274	0.174	0.686
Health	0.070	0.920	0.944	0.562	0.960	0.480	0.928	–	0.232	0.956
Utilities	0.160	0.922	0.900	0.826	0.952	0.822	0.908	0.918	–	0.908
Others	0.002	0.940	0.268	0.198	0.962	0.248	0.264	0.164	0.074	–