Online Appendix for Ranking Intersecting Distribution Functions

Rolf Aaberge* Tarjei Havnes† Magne Mogstad‡

*Statistics Norway. Corresponding author. Tel: +47 21094864. Fax: +47 22855035. Email: roa@ssb.no.
†University of Oslo
‡University of Chicago and Statistics Norway

A Appendix

A.1 Proofs of Theorems 2.1-2.4

Lemma A.1. *Let H be the family of bounded, continuous and non-negative functions on* $[0,1]$ *which are positive on* $(0,1)$ *and let g be an arbitrary bounded and continuous function on* $[0,1]$. *Then*

$$\int g(t)\,h(t)\,dt > 0 \quad \text{for all } h \in H$$

implies

$$g(t) \geq 0 \quad \text{for all } t \in [0,1]$$

and the inequality holds strictly for at least one $t \in (0,1)$.

Proof. The proof of Lemma A.1 is known from mathematical text books. □

Proof. **Theorem 2.4.**[1] Using integration by parts and inserting $\Lambda_F^2(u)$ and $\Lambda_F^3(u)$ from Equations (2.10) and (2.11), we get that

$$\begin{aligned} W_P(F_1) - W_P(F_0) &= -P''(1)\int_0^1 \left(\Lambda_{F_1}^2(t) - \Lambda_{F_0}^2(t)\right)dt + \int_0^1 P'''(u)\int_0^u \left(\Lambda_{F_1}^2(t) - \Lambda_{F_0}^2(t)\right)dt\,du \\ &= -P''(1)\left(\Lambda_{F_1}^3(1) - \Lambda_{F_0}^3(1)\right) + \int_0^1 P'''(u)\left(\Lambda_{F_1}^3(u) - \Lambda_{F_0}^3(u)\right)du \end{aligned}$$

To prove the equivalence between (i) and (ii), note that if (i) holds then $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}_3$. To prove the converse statement, we restrict to preference functions $P \in \mathscr{P}_3$, for which $P''(1) = 0$. Hence,

$$W_P(F_1) - W_P(F_0) = \int_0^1 P'''(u)\left(\Lambda_{F_1}^3(u) - \Lambda_{F_0}^3(u)\right)du > 0$$

and the desired result is obtained by applying Lemma A.1.

To prove the equivalence between (ii) and (iii), consider a case where we transfer a small amount γ from persons with incomes $F^{-1}(s+h_1)$ and $F^{-1}(t+h_1)$ to persons with incomes $F^{-1}(s)$ and $F^{-1}(t)$, respectively, where $t > s$. Then W_P obeys first-degree DPTS if and only if $P'(s) - P'(s+h_1) > P'(t) - P'(t+h_1)$ which for small h_1 is equivalent to $P''(t) - P''(s) > 0$. Next, we find that, for $t - s$ small, this is equivalent to $P'''(s) > 0$. □

Proof. **Theorem 2.2.** The proof is analogous to the proof of Theorem 2.1, and is based on the

[1] The proof of the equivalence between (i) and (ii) in Theorem 2.1 is analogous to the proof for stochastic dominance in Hadar and Russell (1969) but is included for the sake of completeness.

expression

$$\begin{aligned} W_P(F_1) - W_P(F_0) &= -P''(0)\int_0^1 \left(\Lambda_{F_1}^2(t) - \Lambda_{F_0}^2(t)\right)dt + \int_0^1 P'''(u)\int_u^1 \left(\Lambda_{F_1}^2(t) - \Lambda_{F_0}^2(t)\right)dtdu \\ &= -P''(0)\left(\tilde{\Lambda}_{F_1}^3(0) - \tilde{\Lambda}_{F_0}^3(0)\right) - \int_0^1 P'''(u)\left(\tilde{\Lambda}_{F_1}^3(u) - \tilde{\Lambda}_{F_0}^3(u)\right)du \end{aligned}$$

which is obtained by using integration by parts and inserting $\tilde{\Lambda}_F^3(u)$ defined by Equation (2.16). Thus, by arguments like those in the proof of Theorem 2.1 the results of Theorem 2.2 are obtained. □

Proof. **Equivalence between (i) and (ii) in Theorem 2.3.** To examine the case of i^{th}-degree upward inverse stochastic dominance, we integrate $W_P(F_1) - W_P(F_0)$ by parts i times,

$$\begin{aligned} W_P(F_1) - W_P(F_0) &= \sum_{j=2}^{i-1}(-1)^{j-1}P^{(j)}(1)\left[\Lambda_{F_1}^{j+1}(1) - \Lambda_{F_0}^{j+1}(1)\right] \\ &\quad + (-1)^{i-1}\int_0^1 P^{(i)}(u)\left[\Lambda_{F_1}^i(u) - \Lambda_{F_0}^i(u)\right]du \end{aligned} \tag{A.1}$$

and use this expression in constructing the proof of the equivalence between (i) and (ii).

Assume first that (i) in Theorem 2.3 is true, i.e. $\Lambda_{F_1}^i(u) - \Lambda_{F_0}^i(u) \geq 0$ for all $u \in [0,1]$ and $>$ holds for at least one $u \in (0,1)$. Then $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}_i$.

Conversely, assume that $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}_i$. For this family of social welfare functions, we have that

$$W_P(F_1) - W_P(F_0) = (-1)^{i-1}\int_0^1 P^{(i)}(u)\left(\Lambda_{F_1}^i(u) - \Lambda_{F_0}^i(u)\right)du > 0$$

Then, as demonstrated by Lemma A.1, the desired result can be obtained by a suitable choice of $P \in \mathscr{P}_i$. □

Proof. **Equivalence between (ii) and (iii) in Theorem 2.3.** We prove the equivalence between (ii) and (iii) in Theorem 2.3 by using mathematical induction. To this end it is convenient to introduce the following notation. Let H_1, H_2 and H_{j+1} be defined by

$$H_1(v,h_1) = P'(v) - P'(v+h_1) \tag{A.2}$$

$$H_2(s,t,h_1) = H_1(s,h_1) - H_1(t,h_1) \tag{A.3}$$

$$H_{j+1}(s,t,h_1,h_2,\ldots,h_j) = H_j(s,t,h_1,h_2,\ldots,h_{j-1}) \tag{A.4}$$

$$-H_j(s+h_j,t+h_j,h_1,h_2,\ldots,h_{j-1}) \quad \text{for } j = 2,3,\ldots \tag{A.5}$$

Moreover, let

$$H_2^{(1)}(s,t) = \lim_{h_1 \to 0}\frac{1}{h_1}H_2(s,t,h_1) \tag{A.6}$$

and

$$H_{j+1}^{(j)}(s,t) = \lim_{h_j\to 0}\cdots\lim_{h_1\to 0}\frac{1}{\prod_{k=1}^{j}h_k}H_{j+1}\left(s,t,h_1,h_2,\ldots,h_j\right) \quad \text{for } j=2,3,\ldots. \tag{A.7}$$

It follows from Theorem 2.1 and the properties of the admissible weigthing functions $P\in\mathscr{P}$ that W_P obeys the Pigou-Dalton principle of transfers and first-degree DPTS if and only if $P''(t)<0$ and $P'''(t)>0$. From Equations (2.27), (2.1) and (A.2)–(A.7), we then get that W_P obeys second-degree DPTS if and only if

$$H_3^{(2)}(s,t) > 0 \quad \text{for } s<t. \tag{A.8}$$

Inserting (A.4), (A.3) and (A.2) for $j=2$ yields

$$\begin{aligned}
H_3^{(2)}(s,t) &= \lim_{h_2\to 0}\lim_{h_1\to 0}\frac{1}{h_1h_2}H_3(s,t,h_1,h_2)\\
&= \lim_{h_2\to 0}\lim_{h_1\to 0}\frac{1}{h_1h_2}\left[H_2(s,t,h_1)-H_2(s+h_2,t+h_2,h_1)\right]\\
&= \lim_{h_2\to 0}\frac{1}{h_2}\left(H_2^{(1)}(s,t)-H_2^{(1)}(s+h_2,t+h_2)\right)\\
&= \lim_{h_2\to 0}\frac{1}{h_2}\lim_{h_1\to 0}\frac{1}{h_1}\left\{P'(s)-P'(s+h_1)-\left(P'(t)-P'(t+h_1)\right)\right.\\
&\qquad \left.-\left[P'(s+h_2)-P'(s+h_1+h_2)-\left(P'(t+h_2)-P'(t+h_1+h_2)\right)\right]\right\}\\
&= \lim_{h_2\to 0}\frac{1}{h_2}\left[-P''(s)+P''(s+h_2)-\left(P''(t)+P''(t+h_2)\right)\right] = P^{(3)}(s)-P^{(3)}(t).
\end{aligned}$$

Inserting $t=s+h$, we find, for small h, that this is equivalent to $P^{(4)}(s)<0$.

Next, assume that

$$H_j^{(j-1)}(s,t) = (-1)^{j-1}\left(P^{(j)}(s)-P^{(j)}(t)\right). \tag{A.9}$$

It follows from Theorem 2.1 and the above that (A.9) is true for $j=2$ and $j=3$. Inserting (A.4) in (A.7), we get

$$\begin{aligned}
H_{j+1}^{(j)}(s,t) &= \lim_{h_j\to 0}\cdots\lim_{h_1\to 0}\frac{1}{\prod_{k=1}^{j}h_k}\left(H_j\left(s,t,h_1,h_2,\ldots,h_{j-1}\right)-H_j\left(s+h_j,t+h_j,h_1,h_2,\ldots,h_{j-1}\right)\right)\\
&= \lim_{h_j\to 0}\cdots\lim_{h_2\to 0}\frac{1}{\prod_{k=2}^{j}h_k}\left(H_j^{(1)}\left(s,t,h_1,h_2,\ldots,h_{j-1}\right)-H_j^{(1)}\left(s+h_j,t+h_j,h_1,h_2,\ldots,h_{j-1}\right)\right)\\
&= \lim_{h_j\to 0}\frac{1}{h_j}\left(H_j^{(j-1)}(s,t)-H_j^{(j-1)}\left(s+h_j,t+h_j\right)\right),
\end{aligned}$$

which by inserting (A.9) yields

$$H_{j+1}^{(j)}(s,t) = (-1)^j\left(P^{(j+1)}(s)-P^{(j+1)}(t)\right).$$

Thus, (A.9) is proved to be true by induction.

Since W_P defined by Equation (**??**) obeys the $(i-1)$th-degree DPTS if and only if $H_i^{(i-1)}(s,t) > 0$ for $s < t$, we get from (A.9) that this condition is equivalent to $(-1)P^{(i+1)}(s) > 0$. □

Proof. **Theorem 2.4.** The proof follows exactly the reasoning used in the proof of Theorem 2.3, using the following expression,

$$W_P(F_1) - W_P(F_0) = -\sum_{j=2}^{i-1} P^{(j)}(0)\left[\tilde{\Lambda}_{F_1}^{j+1}(0) - \tilde{\Lambda}_{F_0}^{j+1}(0)\right] - \int_0^1 P^{(i)}(u)\left[\tilde{\Lambda}_{F_1}^{i}(u) - \tilde{\Lambda}_{F_0}^{i}(u)\right] du$$

which is obtained by using integration by parts i times. □

A.2 Appendix: Asymptotic theory

This section develops distribution theory to test for upward and downward inverse stochastic dominance of any degree.[2]

Let X be an income variable with cumulative distribution function F and mean μ. Let $[a,b]$ be the domain of F where F^{-1} is the left inverse of F and $F^{-1}(0) \equiv a \geq 0$. Let $X_1, X_2, \ldots, X_n$ be independent random variables with common distribution function F and let F_n be the corresponding empirical distribution function.

Estimation of dominance functions

Since the parametric form of F is not known, it is natural to use the empirical distribution function F_n to estimate F and to use

$$\Lambda^i_{F_n}(u) = \frac{1}{(i-2)!} \int_0^u (u-t)^{i-2} F_n^{-1}(t)\, dt, \quad 0 \leq u \leq 1, i = 2, 3, \ldots$$

to estimate $\Lambda^i_F(u)$, where F_n^{-1} is the left inverse of F_n, and to use

$$\tilde{\Lambda}^i_{F_n}(u) = \frac{1}{(i-2)!} \left[(1-u)^{i-2} \int_0^1 F_n^{-1}(t)\, dt - \int_u^1 (t-u)^{i-2} F_n^{-1}(t)\, dt \right], \quad 0 \leq u \leq 1, i = 3, 4, \ldots$$

to estimate $\tilde{\Lambda}^i_F(u)$.

To obtain explicit expressions for $\Lambda^i_{F_n}(u)$ and $\tilde{\Lambda}^i_{F_n}(u)$, let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the ordered $X_1, X_2, \ldots, X_n$ and $\bar{X} = \frac{1}{n}\sum_{k=1}^n X_k$. For $u = k/n$, we have

$$\Lambda^i_{F_n}\left(\frac{k}{n}\right) = \frac{1}{(i-2)!} \frac{1}{n} \sum_{j=1}^{k} \left(\frac{k-j}{n}\right)^{i-2} X_{(j)}, \quad k = 1, 2, \ldots, n$$

and

$$\tilde{\Lambda}^i_{F_n}\left(\frac{k}{n}\right) = \frac{1}{(i-2)!} \left[\left(1 - \frac{k}{n}\right)^{i-2} \bar{X} - \frac{1}{n} \sum_{j=k}^{n} \left(\frac{j-k}{n}\right)^{i-2} X_{(j)} \right], \quad k = 1, 2, \ldots, n.$$

Since F_n is a consistent estimator of F, $\Lambda^i_{F_n}(u)$ and $\tilde{\Lambda}^i_{F_n}(u)$ are consistent estimators of $\Lambda^i_F(u)$ and $\tilde{\Lambda}^i_F(u)$.

[2] Andreoli (2018) develops a test with pointwise confidence intervals which essentially implies multiple testing across quantiles. In contrast, we represent the quantile function as a Gaussian continuous process which allows us to develop confidence bands. For alternative approaches to testing for standard stochastic dominance, see Abadie (2002), Anderson (1996), Barrett and Donald (2003), Linton et al. (2005), and Davidson and Duclos (2000), among others.

Asymptotic distribution theory

Let the empirical process $Q_n(u)$ be defined by

$$Q_n(u) = \sqrt{n}\left(F_n^{-1}(u) - F^{-1}(u)\right) \tag{A.10}$$

Approximations to the variances of $\Lambda_{F_n}^i(u)$ and $\tilde{\Lambda}_{F_n}^i(u)$ and the asymptotic properties of $\Lambda_{F_n}^i(u)$ and $\tilde{\Lambda}_{F_n}^i(u)$ can be obtained by considering the limiting distribution of the empirical processes $Y_n^i(u)$ and $\tilde{Y}_n^i(u)$ defined by

$$Y_n^i(u) = \sqrt{n}\left[\Lambda_{F_n}^i(u) - \Lambda_F^i(u)\right] = \frac{1}{(i-2)!}\int_0^u (u-t)^{i-2} Q_n(t)\,dt \tag{A.11}$$

and

$$\tilde{Y}_n^i(u) = \sqrt{n}\left[\tilde{\Lambda}_{F_n}^i(u) - \tilde{\Lambda}_F^i(u)\right] = \frac{1}{(i-2)!}\left[(1-u)^{i-2}\int_0^1 Q_n(t)\,dt - \int_u^1 (t-u)^{i-2} Q_n(t)\,dt\right] \tag{A.12}$$

Let $w(u,t)$ be a function of u and t such that $0 \leq w(u,t) \leq 1$ for all $u,t \in [0,1]$ and let $a(u)$ and $b(u)$ be functions of u such that $0 \leq a(u) < b(u) \leq 1$. In order to study the asymptotic behavior of (A.11) and (A.12) it is convenient to consider the empirical process

$$V_n(u) = \int_{a(u)}^{b(u)} w(u,t)\, Q_n(t)\,dt \tag{A.13}$$

which suggests that it will be useful to start with the process $Q_n(u)$ defined in (A.10).

The processes $Q_n(u)$ and $V_n(u)$ are members of the space D of functions on $[0,1]$ which are right-continuous and have left-hand limits. On this space, we use the Skorokhod topology and the associated σ-field (e.g. Billingsley, 1968, p. 111). We let $W_0(t)$ denote a Brownian bridge on $[0,1]$, that is, a Gaussian process with mean zero and covariance function $s(1-t)$, where $0 \leq s \leq t \leq 1$.

Theorem A.1. *Suppose that F has a continuous nonzero derivative f on $[a,b]$. Then $V_n(u)$ converges in distribution to the process*

$$V(u) = \int_{a(u)}^{b(u)} w(u,t)\frac{W_0(t)}{f(F^{-1}(t))}dt$$

Proof. It follows directly from Theorem 4.1 of Doksum (1974) that the empirical process $Q_n(t)$ converges in distribution to the Gaussian Process $W_0(t)/f\left(F^{-1}(t)\right)$. Using the arguments of Durbin (1973, Section 4.4), it follows that $V_n(u)$ as function of $\left(W_0(t)/f\left(F^{-1}(t)\right)\right)$ is continuous in the Skorokhod topology. The results then follow from Billingsley (1968, Th. 5.1). □

The following result states that $V(u)$ is a Gaussian process and thus that $V_n(u)$ is asymptotically normally distributed, both when considered as a process, and for fixed u.

Theorem A.2. *Suppose the conditions of Theorem A.1 are satisfied. Then the process $V(u)$ has the same probability distribution as the Gaussian process*

$$\sum_{j=1}^{\infty} d_j(u) Z_j$$

where $d_j(u)$ is given by

$$d_j(u) = \frac{\sqrt{2}}{j\pi} \int_{a(u)}^{b(u)} w(u,t) \frac{\sin(j\pi t)}{f(F^{-1}(t))} dt$$

and $Z_1, Z_{2,} \ldots$ are independent $N(0,1)$-variables.

Proof. Let

$$Q_N^*(t) = \frac{\sqrt{2}}{f(F^{-1}(t))} \sum_{j=1}^{N} \frac{\sin(j\pi t)}{j\pi} Z_j$$

and note that

$$2\sum_{j=1}^{N} \frac{\sin(j\pi s)\sin(j\pi t)}{(j\pi)^2} = s(1-t) \tag{A.14}$$

Thus, the process $Q_N^*(t)$ is Gaussian with mean zero and covariance function

$$\operatorname{cov}(Q_N^*(s), Q_N^*(t)) = \frac{2}{f(F^{-1}(s)) f(F^{-1}(t))} \sum_{j=1}^{N} \frac{\sin(j\pi s)\sin(j\pi t)}{(j\pi)^2} \longrightarrow \operatorname{cov}(Q(s), Q(t))$$

where

$$Q(t) = \frac{W_0(t)}{f(F^{-1}(t))}$$

In order to prove that Q_N^* converges in distribution to the Gaussian process $Q(t)$, it is, according to Hájek and Šidák (1967, Ths. 3.1.a, 3.1.b and 3.2), enough to show that

$$E[Q_N^*(t) - Q_N^*(s)]^4 \leq M(t-s)^2, \quad 0 \leq s,t, \leq 1$$

where the constant M is independent of N.

Since for normally distributed random variables with mean 0,

$$EX^4 = 3\left[EX^2\right]^2$$

we have

$$
\begin{aligned}
&E\left[Q_N^*(t)-Q_N^*(s)\right]^4 = 3\left[\operatorname{var}\left(Q_N^*(t)-Q_N^*(s)\right)\right]^2 \\
&= 3\left\{2\cdot \operatorname{var}\left[\sum_{j=1}^{N}\frac{1}{j\pi}\left(\frac{\sin(j\pi t)}{f(F^{-1}(t))}-\frac{\sin(j\pi s)}{f(F^{-1}(s))}\right)Z_j\right]\right\}^2 \\
&= 3\left\{2\cdot \sum_{\mathrm{j=1}}^{\mathrm{N}}\left[\frac{1}{j\pi}\left(\frac{\sin(j\pi t)}{f(F^{-1}(t))}-\frac{\sin(j\pi s)}{f(F^{-1}(s))}\right)Z_j\right]^2\right\}^2 \\
&\leq 3\left\{2\cdot \sum_{\mathrm{j=1}}^{\infty}\left[\frac{1}{j\pi}\left(\frac{\sin(j\pi t)}{f(F^{-1}(t))}-\frac{\sin(j\pi s)}{f(F^{-1}(s))}\right)Z_j\right]^2\right\}^2 \\
&= 3\left\{\frac{t(1-t)}{f^2(F^{-1}(t))}+\frac{s(1-s)}{f^2(F^{-1}(s))}-2\frac{\operatorname{cov}(W_0(s),W_0(t))}{f(F^{-1}(s))\,f(F^{-1}(t))}\right\}^2.
\end{aligned}
$$

Since $0 < f(x) < \infty$ on $[a,b]$, there exists a constant $M \geq 0$ such that

$$
f\left(F^{-1}(t)\right) \geq M^{-\frac{1}{4}} \text{ for all } t \in [0,1]
$$

Hence, $Q_N^*(t)$ converges in distribution to the process $Q(t)$. Thus, since $w(u,t)$ is bounded it follows according to Billingsley (1968, Th. 5.1) that

$$
\int_{a(u)}^{b(u)} w(u,t)\,Q_N^*(t)\,dt = \sum_{j=1}^{N} d_j(u)\,Z_j
$$

converges in distribution to the process

$$
\int_{a(u)}^{b(u)} w(u,t)\,P(t)\,dt = \int_{a(u)}^{b(u)} w(u,t)\,\frac{W_0(t)}{f(F^{-1}(t))}dt = Z(u)
$$

□

The following result is obtained from Theorems A.1 and A.2 by inserting $a(u) = 0$, $b(u) = u$ and $w(u,t) = (u-t)^{i-2}/(i-2)!$ in expression (4.4).

Corollary A.1. *Suppose that F has a continuous nonzero derivative f on $[a,b]$. Then $Y_n^i(u)$ converges in distribution to the process*

$$
Y^i(u) = \frac{1}{(i-2)!}\int_0^u (u-t)^{i-2}\,\frac{W_0(t)}{f(F^{-1}(t))}dt
$$

which has the same probability distribution as the Gaussian process

$$
\sum_{j=1}^{\infty} h_j^i(u)\,Z_j
$$

where $h_j^i(u)$ is given by

$$h_j^i(u) = \frac{1}{(i-2)!}\left[\frac{\sqrt{2}}{j\pi}\int_0^u (u-t)^{i-2}\frac{\sin(j\pi t)}{f(F^{-1}(t))}dt\right]$$

and $Z_1, Z_2, \ldots$ are independent $N(0,1)$-variables.

The following result states that $\tilde{Y}_n^i(u)$ converges to a Gaussian process and thus that $\tilde{Y}_n^i(u)$ is asymptotically normally distributed.

Corollary A.2. *Suppose that F has a continuous nonzero derivative f on $[a,b]$. Then $\tilde{Y}_n^i(u)$ converges in distribution to the process*

$$\tilde{Y}^i(u) = \frac{1}{(i-2)!}\left[(1-u)^{i-2}\int_0^1 \frac{W_0(t)}{f(F^{-1}(t))}dt - \int_u^1 (t-u)^{i-2}\frac{W_0(t)}{f(F^{-1}(t))}dt\right]$$

which has the same probability distribution as the Gaussian process

$$\sum_{j=1}^{\infty} \tilde{h}_j^i(u) Z_j$$

where $\tilde{h}_j^i(u)$ is given by

$$\tilde{h}_j^i(u) = \frac{1}{(i-2)!}\frac{\sqrt{2}}{j\pi}\left[(1-u)^{i-2}\int_0^1 \frac{\sin(j\pi t)}{f(F^{-1}(t))}dt - \int_u^1 (t-u)^{i-2}\frac{\sin(j\pi t)}{f(F^{-1}(t))}dt\right]$$

and $Z_1, Z_2, \ldots$ are independent $N(0,1)$-variables.

Proof. Theorem A.1 implies that the process $\tilde{Y}_n^i(u)$ converges in distribution to the process $\tilde{Y}^i(u)$. By inserting for respectively $a(u)=0$, $b(u)=1$ and $w(u,t)=(1-u)^{i-2}/(i-2)!$, and for $a(u)=u$, $b(u)=1$ and $w(u,t)=(t-u)^{i-2}/(i-2)!$ in expression (A.13), it follows from Theorem A.2 that the first term $(1-u)^{i-2}\int_0^1 Q_n(t)\,dt$ of expression (A.12) converges to a process that has the same distribution as $\sum_{j=1}^{\infty}\frac{\sqrt{2}}{j\pi}\left[(1-u)^{i-2}\int_0^1 \frac{\sin(j\pi t)}{f(F^{-1}(t))}dt\right]Z_j$, while the second term $\left[\int_u^1 (t-u)^{i-2} Q_n(t)\,dt\right]$ of expression (A.12) converges to a process that has the same distribution as $\sum_{j=1}^{\infty}\frac{\sqrt{2}}{j\pi}\left[\int_u^1 (t-u)^{i-2}\frac{\sin(j\pi t)}{f(F^{-1}(t))}dt\right]Z_j$. □

By applying Fubini's theorem (e.g. Royden, 1963) and the identity

$$2\sum_{j=1}^{\infty}\frac{sin(j\pi s)\,sin(j\pi t)}{(j\pi)^2} = s(1-t),\ 0 \le s \le t \le 1, \tag{A.15}$$

we get as an immediate consequence of Corollary A.1 the following result.

Corollary A.3. *Under the conditions of Theorem A.1, $Y_n^i(u)$ has asymptotic covariance function given by*

$$v_i^2(u,v) = \sum_{j=1}^{\infty} h_j^i(u)\, h_j^i(v) \tag{A.16}$$

$$= \frac{1}{[(i-2)!]^2}\left\{2\int_{F^{-1}(0)}^{F^{-1}(u)}\int_{F^{-1}(0)}^{y}[(u-F(x))(v-F(y))]^{i-2}F(x)(1-F(y))\,dxdy \right.$$

$$\left. + \int_{F^{-1}(u)}^{F^{-1}(v)}\int_{F^{-1}(0)}^{F^{-1}(u)}[(u-F(x))(v-F(y))]^{i-2}F(x)(1-F(y))\,dxdy\right\}$$

In order to derive the asymptotic covariance function of $\tilde{Y}_n^i(u)$ it proves convenient to introduce the following notation.

$$\lambda_{ikr}(u,v) = \frac{1}{[(i-2)!]^2}\int_{F^{-1}(v)}^{F^{-1}(1)}\int_{F^{-1}(v)}^{y}(F(x)-u)^{k-2}(F(y)-v)^{r-2}F(x)(1-F(y))\,dxdy,$$

$$\gamma_{ikr}(u,v) = \frac{1}{[(i-2)!]^2}\int_{F^{-1}(v)}^{F^{-1}(1)}\int_{F^{-1}(u)}^{F^{-1}(v)}(F(x)-u)^{k-2}(F(y)-v)^{r-2}F(x)(1-F(y))\,dxdy$$

and

$$\tilde{\lambda}_{ikr}(u,v) = \frac{1}{[(i-2)!]^2}\int_{F^{-1}(v)}^{F^{-1}(1)}\int_{F^{-1}(v)}^{y}(F(x)-v)^{k-2}(F(y)-u)^{r-2}F(x)(1-F(y))\,dxdy.$$

Now, similarly as for Corollary A.3, we get the following result from Corollary A.2 by applying Fubini's theorem (e.g. Royden, 1963) and the identity (A.15).

Corollary A.4. *Under the conditions of Theorem A.1, $\tilde{Y}_n^i(u)$ has asymptotic covariance function given by*

$$\eta_i^2(u,v) = \sum_{j=1}^{\infty} \tilde{h}_j^i(u)\,\tilde{h}_j^i(v) \tag{A.17}$$

$$= 2[(1-u)(1-v)]^{i-2}\lambda_{i22}(0,0) - (1-u)^{i-2}[\lambda_{i2i}(u,v)+\lambda_{ii2}(u,v)+\gamma_{i2i}(0,v)]$$

$$-(1-v)^{i-2}[\lambda_{i2i}(u,u)+\lambda_{ii2}(u,u)+\gamma_{i2i}(0,u)] + \left[\lambda_{iii}(u,v)+\tilde{\lambda}_{iii}(u,v)+\gamma_{iii}(u,v)\right]$$

In order to construct confidence intervals for $\Lambda_F^i(u)$ and $\tilde{\Lambda}_F^i(u)$ at fixed points, we apply the results of Theorem A.1 and Corollary A.2, which imply that the distribution of

$$\sqrt{n}\frac{\Lambda_{F_n}^i(u)-\Lambda_F^i(u)}{v_i(u,u)}$$

tends to the $N(0,1)$-distribution for fixed u, where $v_i^2(u,u)$ is given by (A.16), and the distribution of

$$\sqrt{n}\frac{\tilde{\Lambda}^i_{F_n}(u)-\tilde{\Lambda}^i_F(u)}{\eta_i(u,u)}$$

tends to the $N(0,1)$-distribution for fixed u, where $\eta_i^2(u,u)$ is given by (A.17).

Confidence intervals and bands

To get an idea of how reliable $\Lambda^i_{F_n}(u)$ and $\tilde{\Lambda}^i_{F_n}(u)$ are as estimates of $\Lambda^i_F(u)$ and $\tilde{\Lambda}^i_F(u)$, we have to construct confidence bands based on $\Lambda^i_{F_n}(u)$ and $\tilde{\Lambda}^i_{F_n}(u)$, respectively. Such confidence bands can be obtained from statistics of the type

$$K_n=\sqrt{n}\sup\frac{|V_n(u)-V(u)|}{\psi(V_n(u))}$$

where ψ is a continuous nonnegative weight function. By applying Theorems A.1 and A.2 and Billingsley (1968, Th. 5.1), we find that K_n converges in distribution to

$$K=\sup_{0\le u\le 1}\left|\sum_{j=1}^{\infty}\frac{d_j(u)}{\psi(V(u))}Z_j\right|$$

We use the following notation.

$$\begin{aligned}
T_m(u) &= \sum_{j=1}^{m}\frac{d_j(u)}{\psi(V(u))}Z_j\\
T(u) &= \sum_{j=1}^{\infty}\frac{d_j(u)}{\psi(V(u))}Z_j\\
K_m^* &= \sup_{0\le u\le 1}|T_m(u)|
\end{aligned}$$

Since T_m converges in distribution to T, we find by applying Billingsley (1968, Th. 5.1) that K_m^* converges in distribution to K. Hence, for a suitable choice of m and ψ, for instance $\psi=1$, simulation methods may be used to obtain the distribution of K_m^* and thus an approximation for the distribution of K.

Implementation

Following Sverdrup (1976), consider two distributions, F_1 and F_0. We are interested in drawing inference about whether one of the distributions inverse stochastic dominates the other of a given degree. For example, we want to draw statistical inference about upward inverse stochastic dominance of i-degree, i.e. deciding whether $\Lambda^i_{F_1}(u)\ge\Lambda^i_{F_0}(u)$ for all $u\in[0,1]$ *or* $\Lambda^i_{F_1}(u)\le\Lambda^i_{F_0}(u)$ for all $u\in[0,1]$ and the inequalities holds strictly for some $u\in(0,1)$. As explained in Sverdrup (1976), such a testing problem gives three possible decisions. One possibility is that F_1 dominates F_0 of ith-degree, i.e.

$$B_1 = \left\{\Lambda^i_{F_1}(u) \geq \Lambda^i_{F_0}(u) \text{ for all } u \in [0,1] \text{ and the inequality holds strictly for some } u \in (0,1)\right\}$$

while another possibility is that F_0 dominates F_1 of ith-degree, i.e.

$$B_2 = \left\{\Lambda^i_{F_1}(u) \leq \Lambda^i_{F_0}(u) \text{ for all } u \in [0,1] \text{ and the inequality holds strictly for some } u \in (0,1)\right\}$$

and the last possibility is that we cannot draw inference about dominance of ith degree, i.e.

$$B_3 = \{\text{We cannot conclude that either } B_1 \text{ or } B_2 \text{ is true }\}$$

Thus, we have a three-decision problem, which can be considered as a special case of Scheffe's method for multiple comparisons.[3]

The statistical inference problem is to decide whether the data justify statement B_1 or B_2. The test has to satisfy the following conditions

$$Pr(B_1) \leq \varepsilon \text{ when } \Lambda^i_{F_1}(u) \leq \Lambda^i_{F_0}(u) \text{ for all } u \in [0,1]$$

$$Pr(B_2) \leq \varepsilon \text{ when } \Lambda^i_{F_1}(u) \geq \Lambda^i_{F_0}(u) \text{ for all } u \in [0,1].$$

The test statistic is derived under the condition $\Lambda^i_{F_1}(u) = \Lambda^i_{F_0}(u)$ for all $u \in [0,1]$. We thus choose this condition as the null-state (null-hypothesis), not because we have any a priori confidence in the truth of it, but because B_1 and B_2 are contrasts relative to the null-state. We conclude B_1 (B_2) if $\Lambda^i_{F_{n1}}(u) \geq \Lambda^i_{F_{n0}}(u)$ ($\Lambda^i_{F_{n1}}(u) \leq \Lambda^i_{F_{n0}}(u)$) for all $u \in [0,1]$ and the inequality holds strictly for some $u \in (0,1)$, and the empirical dominance is statistically significant (given the choice of ε). We refer to Sverdrup (1976) for further details about the testing procedure. In the empirical analysis, we apply the distribution theory to test for upward and downward inverse stochastic dominance. We estimate the degree of dominance of the two empirical distributions under consideration, say F_{1n} and F_{0n}, and calculate the dominance functions $\Lambda^i_{F_n}(u)$ and $\tilde{\Lambda}^i_{F_n}(u)$ for each $u \in [0,1]$.

Simulations

We now examine the finite-sample performance of our procedure for inference with a simulation study. The goal is to determine whether one income distribution inverse stochastic dominates another income distribution of a given degree. For simplicity, we focus on inverse stochastic dominance of second degree.

[3] Note that it is necessary to treat this as a three-decision problem. Testing first whether or not $\Lambda^i_{F_1}(u) \geq \Lambda^i_{F_0}(u)$ for all $u \in [0,1]$ and then whether or not $\Lambda^i_{F_1}(u) \leq \Lambda^i_{F_0}(u)$ for all $u \in [0,1]$ and the inequalities holds strictly for some $u \in (0,1)$ would produce incorrect inference.

Table A.1: Ranking of income distributions by upward and downward dominance

		Upwards dominance													
	YEAR	1994	1995	1996	1997	1998	1999	2000	2001	2005	2006	2007	2008	2009	2010
	1994		3<	3<	2<	100>	100>	100>	100>	1<	1<	1<	1<	2<	1<
Downwards dominance	1995	3>		2<	1<	100>	100>	100>	97>	1<	1<	1<	1<	2<	100>
	1996	4>	2<		100>	54>	49>	65>	50>	94>	97>	1<	100>	97>	58>
	1997	2<	1<	3<		10>	9>	34>	24>	69>	77>	1<	93>	11>	10>
	1998	3<	3<	3<	3<		3>	1<	41>	1<	1<	1<	1<	17<	41<
	1999	3<	3<	3<	3<	3<		1<	54>	1<	1<	1<	1<	5<	24<
	2000	3<	3<	3<	3<	1<	1<		8>	1<	1<	1<	2<	50<	84<
	2001	3<	3<	3<	3<	3<	3<	3<		1<	2<	1<	2<	32<	41<
	2005	1<	1<	3<	3<	1<	1<	1<	1<		100>	2<	40<	93<	1>
	2006	1<	1<	3<	3<	1<	1<	1<	2<	3>		1<	62<	97<	1>
	2007	1<	1<	1<	1<	1<	1<	1<	1<	2<	1<		1>	1>	1>
	2008	1<	1<	3<	3<	1<	1<	2<	2<	3>	3>	1>		100<	1>
	2009	2<	2<	3<	3<	3<	3<	3>	3>	3>	3>	1>	3>		5>
	2010	1<	3<	3<	3<	3<	3<	3>	3>	1>	1>	1>	1>	3<	

Note: This table makes pairwise yearly comparisons of the income distributions over the period 1994-2010. We report the degree and direction of upwards and downwards inverse stochastic dominance. Upwards (downwards) dominance results are reported above (below) the diagonal. We denote by "<" when the later year dominates the earlier year, and by ">" when the earlier year dominates the later year. Dominance degrees are truncated at 100.

We consider three distributions, F_0, F_1, and F_2. These distributions are chosen so that F_1 upward dominates F_0 of second degree, while F_2 upward dominates F_0 and F_1 of third degree. All three distributions are Champernowne distributions, developed to describe the logarithm of income. The Champernowne distribution has a probability density function given by

$$f(x;\alpha,\lambda,x_0) = \frac{k}{cosh(\alpha(x-x_0)+\lambda}$$

where cosh is a hyperbolic function, (α,λ,x_0) are positive parameters and k is a normalizing constant (which depends on the parameters). Below, we graph the inverse cummulative distribution function of the logarithm of income from F_0, F_1, and F_2. The chosen parameter values of each distribution are reported in the footnote of the figure. The Gini coefficient of F_0 is 0.51, while F_1 and F_2 have Gini coefficients of 0.41 and 0.30, respectively.

From each distribution, we draw 1 000 i.i.d. samples of a given size n. For each set of draws, we compare the emirical distributions F_{1n} and F_{0n}, F_{2n} and F_{0n}, and F_{2n} and F_{1n}. For each comparison, we evaluate whether $\Lambda^2_{F_{n1}}(u) > \Lambda^2_{F_{n0}}(u)$ at each percentile of $u \in [0,1]$. We repeat this procedure for each of the 1000 draws, and compute the probability that $\Lambda^2_{F_{n1}}(u)$ exceeds $\Lambda^2_{F_{n0}}(u)$ at a given u. The results from these simulations are summarized in Tables A.2-A.4. In these table, we compare the simulation results to the asymptotic ones. We vary the sample size across the columns.

Consider first Table A.2 where we compare F_{1n} and F_{0n}. The first two columns use a sample size of 6 000, the next two columns use a sample size 10 000, and last two columns use a sample size of 20 000. For each sample size, the asymptotic results are very similar to the results from the simulation. For values of u below 1, we consistently find that the probability that $\Lambda^2_{F_1}(u)$ exceeds $\Lambda^2_{F_0}(u)$ approaches 1. Thus, we focus on u equal to 1. With a sample size of 6 000, we find a 90 percent probability that $\Lambda^2_{F_1}(u)$ exceeds $\Lambda^2_{F_0}(u)$ for u equal to 1. At this value of u, the probability increases to about 95 percent with a sample size of 10 000 and approaches 99 percent with 20 000 observations.

In Tables A.3-A.4, we do the same simulation excercise, except now we consider distributions where one upward dominates the other of third degree but not second degree. Again, the asymptotic results are very similar to the results from the simulation. As sample sizes grow, the probability that $\Lambda^2_{F_2}(u)$ exceed $\Lambda^2_{F_0}(u)$ and the probability that $\Lambda^2_{F_2}(u)$ exceed $\Lambda^2_{F_1}(u)$ are correctly approaching zero for some value(s) of u. By comparison, even with fairly small sample sizes we find a probability close to one for $\Lambda^3_{F_2}(u)$ exceeding $\Lambda^3_{F_0}(u)$ and $\Lambda^3_{F_2}(u)$ exceeding $\Lambda^3_{F_1}(u)$ at each value of $u \in [0,1]$.

Table A.2: Probability that $\Lambda^2_{F_1}(u)$ exceeds $\Lambda^2_{F_0}(u)$ for $u \in [0,1]$

	Sample Size: 6 000		Sample Size: 10 000		Sample Size: 20 000	
	Asymptotic	Simulation	Asymptotic	Simulation	Asymptotic	Simulation
$u = 0.05$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.10$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.15$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.20$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.25$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.30$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.35$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.40$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.45$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.50$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.55$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.60$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.65$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.70$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.75$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.80$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.85$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.90$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.95$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 1$	90.1	90.3	95.0	94.8	99.0	98.9

Note: The simulations are described in the text.

Table A.3: Probability that $\Lambda^2_{F_2}(u)$ exceeds $\Lambda^2_{F_0}(u)$ for $u \in [0,1]$

	Sample Size: 19 100		Sample Size: 31 500		Sample Size: 63 000	
	Asymptotic	Simulation	Asymptotic	Simulation	Asymptotic	Simulation
$u = 0.05$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.10$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.15$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.20$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.25$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.30$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.35$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.40$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.45$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.50$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.55$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.60$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.65$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.70$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.75$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.80$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.85$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.90$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.95$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 1$	10.0	10.2	5.0	4.3	1.0	0.9

Note: The simulations are described in the text.

Table A.4: Probability that $\Lambda^2_{F_2}(u)$ exceeds $\Lambda^2_{F_1}(u)$ for $u \in [0,1]$

	Sample Size: 1 500		Sample Size: 2 500		Sample Size: 5 000	
	Asymptotic	Simulation	Asymptotic	Simulation	Asymptotic	Simulation
$u = 0.05$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.10$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.15$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.20$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.25$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.30$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.35$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.40$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.45$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.50$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.55$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.60$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.65$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.70$	100.0	100.0	100.0	100.0	100.0	100.0
$u = 0.75$	99.4	99.3	100.0	100.0	100.0	100.0
$u = 0.80$	96.0	95.7	98.8	98.7	100.0	100.0
$u = 0.85$	84.1	85.8	90.1	88.7	96.6	96.5
$u = 0.90$	59.3	58.6	61.9	60.4	66.6	65.5
$u = 0.95$	29.7	29.7	24.6	25.1	16.5	14.9
$u = 1$	9.9	9.6	4.8	4.6	0.9	1.3

Note: The simulations are described in the text.

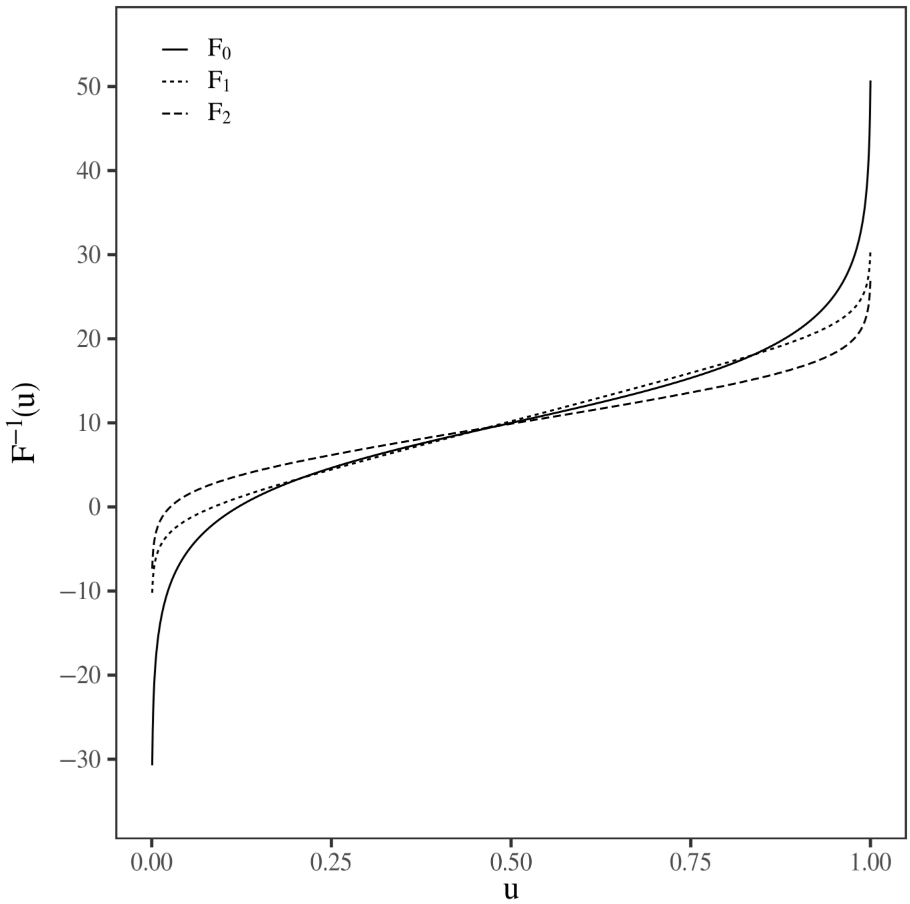

Figure A.1: Inverse cummulative distribution function of the three distributions used in the simulation study

Note: As specified in the text, all three distributions are Champernowne distributions. For F_0 then $\alpha = 0.17, \lambda = 0.1, x_0 = 10$. For F_1 then $\alpha = 0.53, \lambda = 200, x_0 = 10.2$. For F_2 then $\alpha = 0.51, \lambda = 15, x_0 = 9.9$.

References

Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association*, 97:284–292.

Anderson, G. (1996). Nonparametric tests of stochastic dominance in income distributions. *Econometrica*, 64(5):1183–93.

Andreoli, F. (2018). Robust inference for inverse stochastic dominance. *Journal of Business & Economic Statistics*, 36(1):146–159.

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

Billingsley, P. (1968). *Convergence of probability measures*. Wiley series in probability and statistics: Probability and statistics. Wiley.

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

Doksum, K. (1974). Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *The Annals of Statistics*, 2(2):pp. 267–277.

Durbin, J. (1973). *Distribution theory for tests based on the sample distribution function*. Regional conference series in applied mathematics. Society for Industrial and Applied Mathematics.

Hadar, J. and Russell, W. R. (1969). Rules for ordering uncertain prospects. *American Economic Review*, 59(1):25–34.

Hájek, J. and Šidák, Z. (1967). *Theory of rank tests*. Probability and mathematical statistics. Academic Press.

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

Royden, H. (1963). *Real analysis*. Macmillan, New York.

Sverdrup, E. (1976). Significance testing in multiple statistical inference. *Scandinavian Journal of Statistics*, 3(2):73–78.