Working Paper Series

Inference for Inverse Stochastic Dominance

Francesco Andreoli

ECINEQ WP 2013 – 295

ECINEQ 2013 – 295
April 2013

www.ecineq.org

Inference for Inverse Stochastic Dominance*

Francesco Andreoli†
THEMA University of Cergy-Pontoise and University of Verona

Abstract

This note presents an innovative inference procedure for assessing if a pair of distributions can be ordered according to inverse stochastic dominance (ISD). At order 1 and 2, ISD coincides respectively with rank and generalized Lorenz dominance and it selects the preferred distribution by all social evaluation functions that are monotonic and display inequality aversion. At orders higher than the second, ISD is associated with dominance for classes of linear rank dependent evaluation functions. This paper focuses on the class of conditional single parameters Gini social evaluation functions and illustrates that these functions can be linearly decomposed into their empirically tractable influence functions. This approach gives estimators for ISD that are asymptotically normal with a variance-covariance structure which is robust to non-simple randomization sampling schemes, a common case in many surveys used in applied distribution analysis. One of these surveys, the French Labor Force Survey, is selected to test the robustness of Equality of Opportunity evaluations in France through ISD comparisons at order 3. The ISD tests proposed in this paper are operationalized through the user-written "isdtest" Stata routine.

Keywords: Inverse stochastic dominance, inference, influence functions, inequality.

JEL Classification: C12, D31, I32.

* I am grateful to Arnaud Lefranc and Claudio Zoli for valuable suggestions. The usual disclaimer applies.

† **Contact details**: THEMA, Université de Cergy-Pontoise, 33 boulevard du Port, 95011 Cergy CEDEX, France. E-mail : francescondrl@gmail.com.

1 Introduction

This note presents an innovative inference procedure for assessing if a pair of distributions can be ordered according to *inverse stochastic dominance* (ISD), that can be applied to data sampled according to non-simple (i.e. stratified, cluster, or multistage) random sampling schemes (see Zheng 2002). This is often the case for many economic surveys used in applied inequality analysis.

The ISD partial order has been introduced by Muliere and Scarsini (1989) as an alternative to the traditional (direct) stochastic dominance relations (Fishburn 1976, Fishburn and Vickson 1978, Davidson and Duclos 2000). At order one, two distributions can be ordered according to ISD if the *Pen's Parade* chart (i.e. the quantile functions curves) of the dominant distribution lies no point below and at least some points above the one of the dominated distribution. This condition is also called rank dominance (Saposnik 1981). At order two, a similar dominance condition is extended to the integrals of the Pen's parade of the two distributions at any population quantile, that correspond to the *generalized Lorenz curves* (Kolm 1969, Shorrocks 1983).

Generalized Lorenz dominance is extensively adopted as a robust criterion that permits to incorporate efficiency and inequality considerations in evaluating pairs of income distributions. In fact, if a pair of distributions can be ranked according to generalized Lorenz dominance, then all *social evaluation functions* that are increasing and display inequality aversion (i.e. are sensitive to progressive transfers) would agree that the dominant distribution provides higher social welfare than the dominated distribution. Robustness in evaluation cames at a price: when generalized Lorenz curves intersect, the underlying distributions cannot be robustly ordered. A possibility is to use complete orders based on inequality indicators. Alternatively, once can impose stronger dominance conditions. Foster and Shorrocks (1988), Davies and Hoy (1995) and Zoli (1999, 2002) provide conditions for ranking intersecting Generalized Lorenz curves based on third order stochastic dominance or ISD. As pointed out by Muliere and Scarsini (1989), the two criteria do not coincide at order three and above. Davidson and Duclos (2000) study the inference procedures for the stochastic dominance relations at order higher than the second. Here, the focus is on the inference for ISD at higher orders.

Two distributions can be ordered according to ISD at order 3 if, for every quantile, the integral of the generalized Lorenz curve of the dominating distribution is always larger than, or at most equal to the integral of the generalized Lorenz curve of the dominated distribution (see Zoli 1999). The condition can be easily extended to higher orders of dominance making use of recursive integration. As the order of dominance increases the criterion becomes more discriminatory. In fact, ISD at order k induces unanimity in the evaluation of the preferred income distribution within a strict subset of the social evaluation functions displaying inequality aversion, assigning increasing weight to the lowest incomes in the distribution (Zoli 1999, Zoli 2002, Aaberge 2009).

Maccheroni, Muliere and Zoli (2005) (see also Muliere and Scarsini 1989) showed that

ISD at order k can be equivalently verified by comparing *Single Parameter Gini social evaluation functions* (Donaldson and Weymark 1983) parametrized by the order of dominance k. For example, ISD at order 3 is equivalent to verify, for all quantiles, that the well known (absolute) Gini inequality index, computed for all incomes lower than a given income threshold identified by a quantile of the dominating distribution, is always lower than the Gini index computed with respect to the same quantile of the dominated distribution.

The present paper builds on the fundamental relation between ISD and the family of conditional S-Gini social evaluation functions to derive a novel inference procedure to construct ISD comparisons at order higher than the second, denoted the *threshold estimator approach.* The analogy allows to construct empirically tractable and asymptotically valid estimators for the covariance structure between social evaluation functions calculated at different quantiles of any pair of distributions. This is done by (i) linearly decomposing each of the sample estimators of the conditional S-Gini social evaluation functions into the corresponding *influence functions*, that measure the marginal influence of each observation on the estimator (see Barrett and Donald 2009); (ii) showing that the asymptotic variance-covariance structure of the conditional S-Gini social evaluation functions estimators coincides with the variance-covariance structure of their influence functions, that can be rewritten as transformations of the generalized Lorenz curve influence functions; (iii) providing an empirically tractable formulation of the influence functions estimator that can be implemented on the data. Standard asymptotic algebra shows that normality is granted at $\sqrt{n}$ convergence. This result permits to use Wald-type joint test statistics to verify the hypothesis of equality or dominance among pairs of vectors of conditional S-Gini social evaluation functions estimates, that is ISD at any order higher than the second.

The approach presented in this paper can be considered as dual to the approach by Davidson and Duclos (2000), who test if two distributions can be ordered according to *stochastic dominance* making use of a set of linear restrictions on the realizations of the Foster, Greer and Thorbecke (1984) family of poverty indicators, calculated at different poverty lines.

I define a distribution free approach that applies also to data obtained by non-simple random sampling rules. This is a common feature of many datasets used to carry on empirical distribution analysis, such as PSID, CPS or the French Labor Force Survey (Zheng 2002). Other approaches have been proposed in the literature for testing ISD at order 1 and 2 that are distribution free (Beach and Davidson 1983, Davidson and Duclos 2000) and/or robust to non-simple sampling rules (Zheng 1999, Zheng 2002, Kovacevic and Binder 1997). Aaberge (2006, 2007) has proposed an estimator for the covariance structure between the quantiles of the integrals of the generalized Lorenz curve. However, the empirical counterpart of this estimator seems to be consistent only in the case of data obtained from simple random sampling rules. All these different approaches for testing ISD are implemented in a user-written Stata routine called `isdtes`, made available to the research community.

This paper concludes with an application. Recently, Lefranc, Pistolesi and Trannoy

(2009) have exported the analysis of ISD at order 2 to the study of equality of opportunity for income acquisition in France. According to their model, equality of opportunity is granted whenever the income distribution associated to different types of the population (partitioned according to morally irrelevant characteristics, such as parental background) cannot be ranked according to ISD at order 2. I use the French Labor Force Survey (areolar data) to verify the robustness of equality of opportunity assessments, focusing on comparisons of ISD at order 3. The empirical analysis illustrates how different estimators may lead to diverging conclusions.

The rest of the paper is organized as follows. The ISD model and its relations with social evaluation comparisons are formalized in Section 2. The setting of the analysis is in Section 3 while the main convergence results are reported in Section 4. Empirical implementation and ISD estimators are discussed in Section 5. ISD and equality null hypothesis and test statistics are discussed in Section 6 and implemented in Section 7. Finally, Section 8 concludes.

2 Rank dominance, generalized Lorenz dominance and ISD

Let Y be a random variable with distribution $F(.)$ and *inverse (quantile) distribution function.*[1] $F^{-1}(p)$, for $p \in [0,1]$. Following Gastwirth (1971), the integral function

$$GL(p) \;=\; \int_0^p F^{-1}(t)\,dt$$

defines the *Generalized Lorenz (GL) curve* of the distribution $F(y)$. Successive integrals of the GL curve with respect to p lead to the curve $\Lambda^k(p)$, which is defined recursively by the following relations:

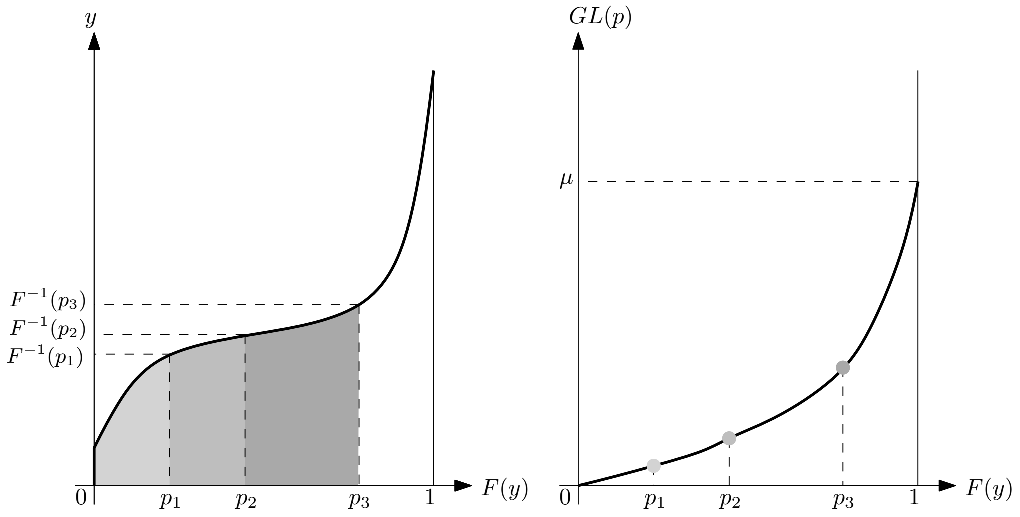
$$\begin{aligned}\Lambda^k(p) &= \int_0^p \Lambda^{k-1}(t)\,dt, \quad p \in [0,1]\\ \Lambda^2(p) &= GL(p).\end{aligned}$$

Muliere and Scarsini (1989) introduced the inverse stochastic dominance partial order as a criterion to rank pairs of distributions. Given two distributions F and G, the relation $F \succcurlyeq_{ISDk} G$ indicates that the distribution F inverse stochastic dominates the distribution G, where ISDk for every $k = 1, \ldots$ is identified as follows:

$$F(y) \;\succcurlyeq_{ISDk}\; G(y) \quad\Leftrightarrow\quad \Lambda^k_F(p) \geq \Lambda^k_G(p) \;\; \forall p \in [0,1].$$

[1] If the cumulative distribution function F (or alternatively G) is only left continuous, we define F^{-1} by the left continuous inverse distribution of F:

$$F^{-1}(p) = \inf\{y \in \mathbb{R}_+ :\; F(y) \geq p\}, \quad \text{with } p \in [0,1].$$



(a) Pen's Parade

(b) GL curve

Figure 1: Quantile function (the *Pen's Parade*), its integrals in three selected coordinates and the generalized Lorenz curve representation.

ISD1 is first order dominance while ISD2 coincides second order stochastic dominance, and they can be verified using the following remark:

Remark 1 For any pair of *cdf*s F and G it holds that:

$$\begin{aligned} F \ \succcurlyeq_{ISD1} \ G \quad &\Leftrightarrow \quad F \text{ rank dominates } \ G, \\ F \ \succcurlyeq_{ISD2} \ G \quad &\Leftrightarrow \quad F \ GL \text{ dominates } \ G. \end{aligned}$$

Figure 1 represents the relation between the *quantile function* of F, also known as the Pen's Parade "of dwarfs and giants" incomes (see Cowell 2000) and its integrals in the space of quantiles, giving the GL curve. ISD1 induces a partial order of distributions. In fact, F rank dominates G if for any position p in the incomes ranking it holds that $F^{-1}(p) \geq G^{-1}(p)$. There are normative considerations that motivate the use od ISD1 in income distributions analysis. In fact, ISD1 is related to agreement within the class $\mathcal{R}^1$ of social evaluation functions (SEF hereafter) admitting the Yaari's (1987) rank dependent model representation. A SEF W maps incomes distributions into social welfare levels, and $W \in \mathcal{R}^1$ only if

$$W(F(y)) \ = \ \int_0^1 w(p)F^{-1}(p)dp,$$

where $w(p)$ is called the *weighting functions* that applies a distortion to the probability p of observing an income lower than $F^{-1}(p)$, such that $w(p) \geq 0 \ \forall p \in [0,1]$ and $\widetilde{w}(p) = \int_0^p w(t)dt \in [0,1]$ with $\widetilde{w}(1) = 1$.

When ISD1 is violated, agreement within $\mathcal{R}^1$ is lost. However, the distributions can be ordered according to weaker criteria than ISD1, such as ISD2. If $F \succcurlyeq_{ISD2} G$ then there is agreement within $\mathcal{R}^2$, the subclass of all SEF in $\mathcal{R}^1$ displaying *inequality aversion*, in ranking F as socially preferred to G. Inequality aversion is modeled as a condition on the weighting function: $w(p)$ must be decreasing in p, i.e. the SEF attach higher weights to lower income realizations.

If ISD2 is violated, agreement within $\mathcal{R}^2$ is lost. Higher orders of ISDk dominance can be used to refine the ISD2 partial order. ISDk is associated with higher order restrictions on the derivatives of the weighting functions $w(p)$, so that if $F \succcurlyeq_{ISDk} G$ then there is agreement within $\mathcal{R}^k$ (a subset of $\mathcal{R}^2$), the class of all SEF displaying higher degrees of inequality aversion, in ranking F as socially preferred to G. The class $\mathcal{R}^k$ can be characterized as follows (Zoli 2002, Maccheroni et al. 2005, Aaberge 2009):

$$\mathcal{R}^k \;=\; \left\{ W \in \mathcal{R} : \; (-1)^{i-1} \cdot \frac{d^i \widetilde{w}(p)}{dp^i} \geq 0, \quad \frac{d^i \widetilde{w}(1)}{dp^i} = 0 \; \forall p \in [0,1] \text{ and } i = 1,2,3,\ldots,k \right\}.$$

It follows that $F \succcurlyeq_{ISDk} G$ if and only if $W(F) \geq W(G)$ $\forall W \in \mathcal{R}^k$. Since ISD$k$ refines ISD1, $F \succcurlyeq_{ISDk} G$ implies $F \succcurlyeq_{ISDl} G$, for all $l > k$ but not the inverse.[2]

Zoli (2002) and Maccheroni et al. (2005) investigate the relation between ISDk and the family of *Generalized Single Parameter Gini (S-Gini) SEF* (Donaldson and Weymark 1983), denoted $W^k(F)$ and defined as follows:[3]

$$W^k(F) \;=\; k \int_0^1 (1-p)^{k-1} F^{-1}(p) dp.$$

The *conditional* Generalized S-Gini SEF at population percentile q, denoted $W^k(q, F)$ (see for instance Zoli 2002) is:

$$W^k(q,F) \;:=\; k \int_0^q (q-p)^{k-1} F^{-1}(p) dp, \tag{1}$$

where $W^k(1,F) = W^k(F)$. Restriction on the parameter k permit to model increasing inequality aversion. The parameter is also related to ISDk:

Remark 2 Given a pair of distributions F and G:

$$F \succcurlyeq_{ISDk} G \quad \Leftrightarrow \quad W^{k-1}(q,F) \geq W^{k-1}(q,G) \text{ for all } q \in [0,1].$$

Therefore, ISD3 can be tested by looking at dominance for all conditional Gini indices $Gini(F)$, since $W^2(F) = \mu_F(1 - Gini(F))$, where μ_F is the average of F. The Remark 2 sets conditions for ISD that play the same role as the Foster et al. (1984) generalized poverty index in assessing direct stochastic dominance at orders higher than the second. The inference procedure studied here is therefore dual, and complements, the analysis by Davidson and Duclos (2000) (Theorem 1).

[2]The rank dependent model differs from the expected utility model for preferences. At order 1 and 2, *stochastic dominance* SD (Fishburn 1976, Fishburn and Vickson 1978) and ISD coincides and the two representations of SEF are equivalent. This does not hold for dominance at order 3 and above.

[3]One can derive from this SEF the whole family of absolute or relative single parameter *S-Gini* inequality measures.

3 Setting

Let $\{Y_i\}_{i=1}^n$ be a sequence of independent random variables Y_i with common distribution F and inverse F^{-1}. The distribution F represents, for instance, an income distribution, where $Y_{(i)}$ indicates that the random variables have been ordered according to the index i, ranging from the lowest $Y_{(1)}$ to the highest $Y_{(n)}$ position. Let assume that a random samples is available, so that income distributions can be estimated from the data. A sample $\mathcal{Y}$ is a collection of n realizations from identically (but not necessarily independently, if not otherwise stated) distributed variables, thus accommodating both simple and non-simple sampling schemes. The realizations of Y_i are denoted with small letters y such that $y_1 \leq \ldots \leq y_i \leq \ldots \leq y_n$. To simplify the exposition, I assume that all the observed values in the sample are distinct (thus n values are observed), so that the index i can be used to denote observations. Moreover, $[y]$ denotes the integer part of the real number y while $\mathbf{1}(.)$ is the indicator function taking value one whenever its argument is verified.

The empirical distribution of F is $\widehat{F}(y) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(Y_i \leq y)$ while the empirical quantile function is $\widehat{F}^{-1}(p) = \inf\{y : \widehat{F}(y) \geq p\}$.

The analysis of ISD permits to compare pairs of distributions F and G using their empirical counterparts $\widehat{F}$ and $\widehat{F}$ respectively, that can be obtained from either paired or unpaired samples (in general $n_F \neq n_G$). I make clear from the outset two important points. First, I verify ISD relations for discrete processes, defined over a finite amount of population quantiles, in order to assess when $F \succcurlyeq_{ISDk} G$ (as in Dardanoni and Forcina 1998, Lefranc et al. 2009), rather than studying continuous processes (as in Barrett and Donald 2003). Second, the results are derived under the assumptions that samples have a simple random structure. I show that the extension to non-simple sampling schemes is immediate under the influence function decomposition.

4 Convergence results for testing ISDk

4.1 Tools and methods

I propose methods for estimating and testing assumptions on F, whose parametric form is in general not known by the econometrician. Following Muliere and Scarsini (1989), one gets the following equivalent representation of the $(k-1)$-th integral of $\widehat{F}^{-1}$:

$$\widehat{\Lambda}^k(p) \quad = \quad \frac{1}{(k-2)!}\int_0^p (p-t)^{k-2}\widehat{F}^{-1}(t)dt, \quad p \in [0,1], \;\; k = 2, 3, \ldots \tag{2}$$

If $\widehat{F}$ is a consistent estimator of F, then $\widehat{\Lambda}^k(p)$ must be a consistent estimator for $\Lambda^k(p)$. This first consideration leads to the first estimation approach, based on the *direct estimator* of ISDk. In practice, the $\Lambda^k(p)$ curve is characterized by a set of ordinates corresponding to a set of pre-chosen m abscissae indexed by $p_1, \ldots, p_m$ so that $0 < p_1 \leq \ldots \leq p_m \leq 1$. The empirical ordinates of the curves are denoted by the sequence $\widehat{\Lambda}^k(p_1), \ldots, \widehat{\Lambda}^k(p_m)$, which in compact vector notation gives:

$$\widehat{\mathbf{\Lambda}}^k = \left(\widehat{\Lambda}^k(p_1), \ldots, \widehat{\Lambda}^k(p_m)\right)^t \in \mathbb{R}^m,$$

with $\mathbf{\Lambda}^k$ being the corresponding vector in the population. The direct approach posits dominance or equality conditions among pairs of vectors $\mathbf{\Lambda}_F^k$ and $\mathbf{\Lambda}_G^k$ associated to distributions F and G respectively, as in Beach and Davidson (1983).

Alternatively, the empirical counterpart of the conditional Generalized Gini SEF in (1) can be obtained by replacing $\widehat{F}^{-1}$ in $W^k(q,F)$ to obtain $\widehat{W}^k(q,\widehat{F})$. If $\widehat{F}$ is a consistent estimator of F, then $\widehat{W}^k(q,\widehat{F})$ is a consistent estimator for $W^k(q,F)$. This consideration motivates the second approach for testing ISDk based on a *threshold estimator*. It requires to compute conditional Generalized Gini SEF for a finite sequence of m thresholds $q_1,\ldots,q_m$ so that $0 < q_1 \leq \ldots \leq q_m \leq 1$. The empirical counterparts of the indicators are denoted by the sequence $\widehat{W}^k(q_1,\widehat{F}),\ldots,\widehat{W}^k(q_m,\widehat{F})$, which in compact vector notation gives:

$$\widehat{\mathbf{W}}_F^k = \left(\widehat{W}^k(q_1,\widehat{F}), \ldots, \widehat{W}^k(q_m,\widehat{F})\right)^t \in \mathbb{R}^m,$$

with $\mathbf{W}_F^k$ being the corresponding vector in the population. The threshold approach posits dominance or equality conditions among pairs of vectors $\mathbf{W}_F^k$ and $\mathbf{W}_G^k$ associated to distributions F and G respectively. The approach has not be yet taken into consideration in the literature and it is, in some sense, dual to the Davidson and Duclos (2000) analysis.

This section investigates the asymptotic distribution and properties of the random vectors $\sqrt{n}\left(\widehat{\mathbf{\Lambda}}^k - \mathbf{\Lambda}^k\right)$ and $\sqrt{n}\left(\widehat{\mathbf{W}}^k - \mathbf{W}^k\right)$ for $k = 3,4,\ldots$. When $k = 1,2$, the asymptotic distribution of the first estimator has been already established in Lemma 1 and Theorem 1 by Beach and Davidson (1983). When $k \geq 3$, some results in Aaberge (2006, 2007) have demonstrated the asymptotic normality of the direct estimator. None of the two approaches, however, can be easily extended to the case in which only non-simple random samples are available. This points highlight the relevance of the threshold estimator.

4.2 The direct estimator approach

For a pair of empirical distributions $\widehat{F}$ and $\widehat{G}$, the *direct approach* requires to compare the estimators $\widehat{\mathbf{\Lambda}}_F^k$ and $\widehat{\mathbf{\Lambda}}_G^k$ at a finite number of population percentiles. The ISDk can be statistically tested making use of joint tests for the equality or dominance null hypothesis. To derive the test distribution, one needs the asymptotic distribution, the asymptotic covariance matrix and its estimator for both vectors of estimates $\widehat{\mathbf{\Lambda}}_F^k$ and $\widehat{\mathbf{\Lambda}}_G^k$.

For a given distribution $\widehat{F}$, the asymptotic distribution and the respective covariance matrix of the quantile functions and the GL curve (that are $\widehat{\mathbf{\Lambda}}^k$ for $k = 1,2$ respectively) have been derived by Beach and Davidson (1983). The main results are summarized in the following Lemma:

Lemma 1 *Suppose that for a set of proportions* $\{p_j | j = 1,\ldots,m\}$ *such that* $0 < p_1 <$

$\ldots < p_m < 1$, $\widehat{\mathbf{\Lambda}} = \left(\widehat{F}^{-1}(p_1), \ldots, \widehat{F}^{-1}(p_m)\right)^t$ is a vector of m sample quantiles and $\widehat{\mathbf{\Lambda}}^2 = \left(\widehat{GL}(p_1), \ldots, \widehat{GL}(p_m)\right)^t$ is a vector of m ordinates of the GL curve estimator (where $\widehat{GL}(1) = \widehat{\mu}$, the sample mean) obtained from a sample of size n drawn from a continuous population density $f(y)$ with cdf $F(y)$ which is strictly monotonic with quantile function $F^{-1}(p)$. Then:

i) *the vector $\sqrt{n}\left(\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\right)$ converges in distribution to a m variate normal distribution with mean zero and asymptotic covariance matrix $\mathbf{\Sigma}^1$, whose element j, j' corresponding to population proportions p_j and $p_{j'}$ is:*

$$\sigma^1(j, j') = \frac{p_j(1 - p_{j'})}{f(F^{-1}(p_j))f(F^{-1}(p_{j'}))}.$$

ii) *the vector $\sqrt{n}\left(\widehat{\mathbf{\Lambda}}^2 - \mathbf{\Lambda}^2\right)$ converges in distribution to a m variate normal distribution with mean zero and asymptotic covariance matrix $\mathbf{\Sigma}^2$, whose element j, j' corresponding to population proportions p_j and $p_{j'}$ is:*

$$\begin{aligned}\sigma^2(j, j') &= p_j v_{p_j}^2 + p_j(1 - p_{j'})(F^{-1}(p_j) - \mu_{p_j})(F^{-1}(p_{j'}) - \mu_{p_{j'}}) + \\ &\quad + p_j(F^{-1}(p_j) - \mu_{p_j})(\mu_{p_{j'}} - \mu_{p_j}) \qquad for\ p_j \leq p_{j'},\end{aligned}$$

where $v_{p_j}^2$ and μ_{p_j} are respectively the variance and expected value of a random variable Y distributed as F conditional on $Y \leq F^{-1}(p_j)$.

The sample counterparts of $\sigma^1(j, j')$ and $\sigma^2(j, j')$ can be obtained by replacing the population moments (quantiles, population shares and the conditional means and variances) with the respective sample estimators, while f is non-parametrically identified.

The result in Lemma 1 can be extended to compare the integrals of the GL curves, that is $\widehat{\mathbf{\Lambda}}^k$ with $k \geq 3$. However, the resulting covariance matrix is hardly empirically tractable.

Proposition 1 *Suppose that for a set of proportions $\{p_j | j = 1, \ldots, m\}$ such that $0 < p_1 < \ldots < p_m < 1$, $\widehat{\mathbf{\Lambda}}^k = \left(\widehat{\Lambda}^k(p_1), \ldots, \widehat{\Lambda}^k(p_m)\right)^t$ for $k = 3, 4, \ldots$ is a vector of m ordinates of the estimator of $\Lambda^k(.)$, the $(k-1)$-th integral of the quantile function, obtained from a sample of size n drawn from a continuous population density $f(y)$ with cdf $F(y)$ which is strictly monotonic with quantile function $F^{-1}(p)$. Then the vector $\sqrt{n}\left(\widehat{\mathbf{\Lambda}}^k - \mathbf{\Lambda}^k\right)$ converges in distribution to a m variate normal distribution with mean zero and asymptotic covariance matrix $\mathbf{\Sigma}^k$ for $k = 3, 4, \ldots$. The element j, j' of $\mathbf{\Sigma}^k$ corresponding to population proportions*

p_j *and* $p_{j'}$ *is:*

$$\begin{aligned}\sigma^k(j,j') &= \frac{1}{(k-2)!}\left[p_j\Lambda^{k-1}(p_j)+(k-1)\Lambda^k(p_j)\right]\left[(1-p_{j'})\Lambda^{k-1}(p_{j'})+(k-1)\Lambda^k(p_{j'})\right]+\\ &+\int_0^{F^{-1}(p_j)}(p_j-F(x))^{k-2}F(x)dx\int_x^{F^{-1}(p_j)}(p_{j'}-F(y))^{k-2}dy \;-\\ &-\int_0^{F^{-1}(p_j)}(p_j-F(x))^{k-2}dx\int_x^{F^{-1}(p_j)}(p_{j'}-F(y))^{k-2}F(y)dy.\end{aligned}$$

Proof. The empirical quantile function estimates only a lower bound for the real population quantile at population share p, that is $\widehat{\Lambda}^k(p) = \frac{1}{n}\Lambda^k(p) + O(n^{-1})$. Using $E[F^{-1}(p)] = F^{-1}(p) + o(n^{-1/2})$ from Lemma 1 (result i)) and the linearity of the operator in (2), one obtains that:

$$E\left[\widehat{\Lambda}^k(p)\right] = \Lambda^k(p) + o(n^{-1/2}).$$

The central limit theorem applies and therefore the vector $\sqrt{n}\left(\widehat{\boldsymbol{\Lambda}}^k - \boldsymbol{\Lambda}^k\right)$ is a multivariate normal with zero means and finite covariance. Consider:

$$\mathrm{cov}\left[n^{-1/2}\widehat{\Lambda}^k(p_j),\ n^{-1/2}\widehat{\Lambda}^k(p_{j'})\right] \quad \text{where } p_j \le p_{j'}. \tag{3}$$

The empirical estimator of $\widehat{\Lambda}^k(p)$ is given by:

$$\widehat{\Lambda}^k(p_j) = \frac{1}{(k-2)!}\sum_{i=1}^{[p_jn]}\left(\frac{[p_jn]-i}{n}\right)^{k-2} Y_{(i)} + o(n^{-1}),$$

so that, for any $p_j \le p_{j'}$, the covariance in (3) can be rewritten as:

$$\frac{1}{n}\frac{1}{(k-2)!}\sum_{i=1}^{[p_jn]}\sum_{h=1}^{[p_{j'}n]}\left(\frac{[p_jn]-i}{n}\right)^{k-2}\left(\frac{[p_{j'}n]-h}{n}\right)^{k-2}\mathrm{cov}(Y_{(i)},Y_{(h)}) + o(n^{-1}). \tag{4}$$

Making use of the consistent estimator of the covariance $\mathrm{cov}\left[Y_{(i)},Y_{(h)}\right]$ in Lemma 1, it is possible to write (4) as:

$$\begin{aligned}&\frac{1}{n}\frac{1}{[(k-2)!]^2}\sum_{i=1}^{[p_jn]}\sum_{h=1}^{i}\left(\frac{[p_jn]-i}{n}\right)^{k-2}\left(\frac{[p_{j'}n]-h}{n}\right)^{k-2}\frac{\left(\frac{h}{n}\right)\left(1-\frac{i}{n}\right)}{n\widehat{F}'(\widehat{F}^{-1}(h/n))\widehat{F}'(\widehat{F}^{-1}(i/n))}+\\ &+\frac{1}{n}\frac{1}{[(k-2)!]^2}\sum_{i=1}^{[p_jn]}\sum_{h=i+1}^{[p_{j'}n]}\left(\frac{[p_jn]-i}{n}\right)^{k-2}\left(\frac{[p_{j'}n]-h}{n}\right)^{k-2}\frac{\left(1-\frac{i}{n}\right)\left(\frac{h}{n}\right)}{n\widehat{F}'(\widehat{F}^{-1}(h/n))\widehat{F}'(\widehat{F}^{-1}(i/n))}+o(n^{-1}).\end{aligned} \tag{5}$$

The estimator $n\,\mathrm{cov}\left[n^{-1/2}\widehat{\Lambda}^k(p_j),\ n^{-1/2}\widehat{\Lambda}^k(p_{j'})\right]$ can be estimated with asymptotic precision equal to $o(1)$. As a consequence, summations can be replaced by integrals to obtain the following formulation of (5):

$$\int_0^{p_j} (p_j - p)^{k-2}(1-p)\frac{dF^{-1}(p)}{dp}\int_0^{p}(p_{j'}-q)^{k-2}q\frac{dF^{-1}(q)}{dq}dqdp + \\ + \int_0^{p_j} (p_j - p)^{k-2}p\frac{dF^{-1}(p)}{dp}\int_p^{p_{j'}}(p_{j'}-q)^{k-2}(1-q)\frac{dF^{-1}(q)}{dq}dqdp.$$

After a change in variables, this integral can be written as:

$$\int_0^{F^{-1}(p_j)} (p_j - F(x))^{k-2}(1-F(x))dx\int_0^{x}(p_{j'}-F(y))^{k-2}F(y)dy + \\ + \int_0^{F^{-1}(p_j)} (p_j - F(x))^{k-2}F(x)dx\int_x^{F^{-1}(p_{j'})}(p_{j'}-F(y))^{k-2}(1-F(y))dy.$$

Integration by parts and the appropriate substitutions of the integration terms give the desired result. ■

A direct consequence of Lemma 1 and Proposition 1 is that for any $k = 1, \ldots$:

$$\widehat{\mathbf{\Lambda}}^k \text{ is asymptotically distributed as } \mathcal{N}\left(\mathbf{\Lambda}^k, \frac{\mathbf{\Sigma}^k}{n}\right) \tag{6}$$

The estimation of $\mathbf{\Sigma}^k$ under the direct approach presents some problems. Either one decides to (non-)parametrically identify f and uses the formula in (5) as an estimator for the variance-covariance matrix, thus incorporating non-standard features of the sampling scheme when estimating f, or one makes use of the Aaberge's (2006) estimator for $\sigma^k(j, j')$, which is valid only under simple random sampling schemes.

4.3 The threshold estimator approach

Making use of the result presented in Remark 2, ISD at order k can be tested by comparing the vectors $\widehat{\mathbf{W}}_F^{k-1}$ and $\widehat{\mathbf{W}}_G^{k-1}$ associated to the distributions F and G, respectively. The asymptotic distribution and the covariance matrix of $\widehat{\mathbf{W}}^k$ studied in this paper are obtained using some results presented in Barrett and Donald (2009), although the procedure for testing ISDk was not considered by the authors. Barrett and Donald study the asymptotic standard errors of the S-Gini SEF and provide its sample estimator making use of the *influence functions* decomposition of $W^k(F)$. I use a similar strategy to decompose the conditional S-Gini SEF $W^k(q, F)$ into their influence functions.

Integrating by parts one can show that the conditional Gini SEF $\widehat{W}^k(p, F)$ can be written as:

$$\widehat{W}^k(p, \widehat{F}) \;=\; k(k-1)\int_0^1 (q-p)^{k-2}\widehat{GL}(p)\mathbf{1}(p \le q)dp \;=\; T(\widehat{H}), \tag{7}$$

where $T(\widehat{H})$ is a scalar valued functional of some process $\widehat{H}$ that is defined on $[0, 1]$. In

(7), $\widehat{H}$ is the GL curve estimator. Using Hadamard differentiability of the functional T (for references, see Barrett and Donald 2009), one obtains that:

$$\sqrt{n}\left(T(\widehat{H}) - T(H)\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} T'_H\left(\phi_i(.,\widehat{H})\right) + o(1),$$

where the random variables $\phi_i(.,\widehat{H})$ are often referred to as the *influence functions*, and give the effect of an observation i on the estimator $\widehat{H}$. The value of the function at q is $\phi(q,\widehat{H})$.

The threshold estimators builds on the fact that the function $T(\widehat{H}) = \widehat{W}^k(p,\widehat{F})$ can be written as a transformation of the $\widehat{GL}$ influence functions, since $\widehat{H} = \widehat{GL}$ and the estimator $\sqrt{n}\left(\widehat{H} - H\right)$ can be decomposed into its influence functions $\phi_i(p,\widehat{GL})$. This point is made clear by the following sequence of passages:

$$\begin{aligned}
\sqrt{n}\left(\widehat{W}^k(q,\widehat{F}) - W^k(q,F)\right) &= k(k-1)\int_0^1 (q-p)^{k-2}\sqrt{n}\left(\widehat{GL}(p) - GL(p)\right)\mathbf{1}(p\leq q)dp \\
&= k(k-1)\int_0^1 (q-p)^{k-2}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\phi_i(p,\widehat{GL})\right)\mathbf{1}(p\leq q)dp + o(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} k(k-1)\int_0^1 (q-p)^{k-2}\phi_i(p,\widehat{GL})\mathbf{1}(p\leq q)dp + o(1) \quad (8) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\phi_i(\widehat{W}^k(q,\widehat{F})) + o(1). \quad (9)
\end{aligned}$$

The terms of the summation in (8) are the influence functions of $\widehat{W}^k(q,\widehat{F})$, denoted $\phi_i(\widehat{W}^k(q,\widehat{F}))$. Extending this result to a $m\times 1$ *vector* of conditional S-Gini SEF computed at different shares q gives:

$$\sqrt{n}\left(\widehat{\boldsymbol{W}}^k - \boldsymbol{W}^k\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\phi_i(\widehat{W}^k(q_1,\widehat{F})),\ldots,\phi_i(\widehat{W}^k(q_m,\widehat{F}))\right)^t + o(1)(1,\ldots,1)^t.$$

The following proposition establishes the asymptotic distribution of $\widehat{W}^k(p,\widehat{F})$.

Proposition 2 *Suppose that for a set of proportions $\{q_j|j=1,\ldots,m\}$ such that $0 < q_1 < \ldots < q_m < 1$, $\widehat{\mathbf{W}}^k = \left(\widehat{W}^k(q_1,\widehat{F}),\ldots,\widehat{W}^k(q_m,\widehat{F})\right)^t$ for $k = 2,3,4,\ldots$ is a vector of m ordinates of the estimator of $W^k(q,F)$, which is the conditional Gini SEF conditional on population share q parametrized at k, obtained from a sample of size n drawn from a continuous population density $f(y)$ with cdf $F(y)$ which is strictly monotonic with quantile function $F^{-1}(p)$. Then, the vector $\sqrt{n}\left(\widehat{\mathbf{W}}^k - \mathbf{W}^k\right)$ converges in distribution to a m variate normal distribution with mean zero and asymptotic covariance matrix $\mathbf{\Sigma}^k$ for $k = 2,3,\ldots,$ The element j,j' of $\mathbf{\Sigma}^k$ corresponding to population proportions q_j and $q_{j'}$ is:*

$$\varsigma^k(q_j,q_{j'}) = E\left[\phi_i(\widehat{W}^k(q_j,\widehat{F}))\,\phi_i(\widehat{W}^k(q_{j'},\widehat{F}))\right]. \quad (10)$$

Proof. By (8), for a given q the random variable $\sqrt{n}\left(\widehat{W}^k(q,\widehat{F}) - W^k(q,\widehat{F})\right)$ satisfies the central limit theorem, and it is therefore asymptotically normal with mean zero and finite variances and covariances. Using (9), $\sqrt{n}\left(\widehat{\mathbf{W}}^k - \mathbf{W}^k\right) = \sum_{i=1}^n \boldsymbol{\phi}_i + o(1)\mathbf{e}_m$ where $\boldsymbol{\phi}_i = \left(\phi_i(\widehat{W}^k(q_1,\widehat{F})),\ldots,\phi_i(\widehat{W}^k(q_m,\widehat{F}))\right)^t$ and $\mathbf{e}_m$ is an $m\times 1$ vector of ones. The random vector is multivariate normal with zero mean and finite covariance equal to

$$\mathrm{E}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \boldsymbol{\phi}_i \cdot \frac{1}{\sqrt{n}}\sum_{i'=1}^n \boldsymbol{\phi}_{i'}^t\right], \tag{11}$$

which, using the properties of the expectations, becomes:

$$\frac{1}{n}\sum_{i=1}^n\sum_{i'=1}^n \mathrm{E}\left[\boldsymbol{\phi}_i \cdot \boldsymbol{\phi}_{i'}^t\right].$$

Since the random vectors $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_{i'}$ are independent by construction of the influence function, one gets that $\mathrm{E}\left[\boldsymbol{\phi}_i \cdot \boldsymbol{\phi}_{i'}^t\right] = 0$ for $i \neq i'$ and that $\mathrm{E}\left[\boldsymbol{\phi}_i \cdot \boldsymbol{\phi}_i^t\right] = \mathrm{E}\left[\boldsymbol{\phi}_{i'} \cdot \boldsymbol{\phi}_{i'}^t\right]$ for all i, i'. This implies that the covariance in (11) is equal to $\mathrm{E}\left[\boldsymbol{\phi}_i \cdot \boldsymbol{\phi}_i^t\right]$, which defines the covariance matrix parameters. The term $\varsigma^k(p_j, p_{j'})$ in (10) corresponds to one of the cells of this matrix. ■

The proposition clarifies that the covariance structure between the conditional S-Gini SEF computed at different quantiles coincides with the influence functions covariance structure. A direct consequence is that for any $k \geq 3$:

$$\widehat{\mathbf{W}}^k \text{ is asymptotically distributed as } \mathcal{N}\left(\mathbf{W}^k,\ \frac{\boldsymbol{\Sigma}^k}{n}\right). \tag{12}$$

The implementation problems and empirical solutions are discussed in the following section. The asymptotic result can be extended to non-simple random samples by incorporating the sampling structure and the weighting scheme in the calculation of (10).

4.4 Bootstrapping techniques

A third possibility for testing ISDk consists in bootstrapping the variance covariance matrix of the threshold estimator of $\widehat{\mathbf{W}}_F^k$. Bootstrapping is often used in the empirical distributional analysis to produce estimators of the standard errors of inequality indicators. Here, I propose to bootstrap vectors of conditional S-Gini SEF for each of the $\{q_1,\ldots,q_m\}$ quantiles separately for a sufficiently large number of sub-samples drawn from the original sample. Then, the covariance matrix of the conditional S-Gini SEF is estimated by the empirical correlation between the indicators values calculated at every bootstrap iteration.

Let $\mathcal{Y}$ be the original sample of size n drawn from the distribution F. Bootstrap computations are constructed conditional on $\mathcal{Y}$. Let a random sample of size n^* drawn with replacement from $\mathcal{Y}$ be denoted as $\mathcal{Y}^b = \{Y_i^b\}_{i=1}^{n^*}$ with empirical distribution $\widehat{F}^b$. It is often the case that $n^* = n$. For every sub-sample $\mathcal{Y}^b$ calculate the conditional S-Gini SEF at order k for a finite number of m quantiles, denoted by $\widehat{W}^k(q_j, \widehat{F}^b)$. By repeatedly drawing random samples from $\mathcal{Y}$, say B times, and calculating for each of the sub-samples the vector of conditional S-Gini SEF, one obtains a $B \times m$ matrix of data. The $m \times m$ empirical covariance matrix computed from these data is the bootstrap estimator of $\mathbf{\Sigma}^k/n$, the asymptotic covariance matrix of $\widehat{\mathbf{W}}^k$ in (12). Hence:

$$\begin{aligned}\frac{\widehat{\varsigma}^k_B(q_j, q_{j'})}{n} &= \frac{1}{B-1}\left(\sum_{b=1}^{B} \widehat{W}^k(q_j, \widehat{F}_b)\widehat{W}^k(q_{j'}, \widehat{F}_b) \;-\; B\overline{\widehat{W}}^k(q_j, \widehat{F}_b)\,\overline{\widehat{W}}^k(q_{j'}, \widehat{F}_b)\right), \quad (13)\\ \overline{\widehat{W}}^k(q, \widehat{F}_b) &= \frac{1}{B}\sum_{b=1}^{B}\widehat{W}^k(q, \widehat{F}_b).\end{aligned}$$

The application of the bootstrap estimator only requires the calculation of a vector of m conditional S-Gini SEF estimators within each re-sampling stage, although in general it does not offer a refinement of the asymptotic approximation illustrated in Proposition 2.

5 Sample implementation

Consider a sample of size n where realizations are $y_1, \ldots, y_n$ ordered such as $y_1 \leq \ldots \leq y_i \leq \ldots \leq y_n$. The estimator of the empirical *cdf* $\widehat{F}$ at any point y is:

$$\widehat{F}(y) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(y_i \leq y), \quad (14)$$

The quantile function estimator at population share p, denoted $\widehat{F}^{-1}(p)$, is:

$$\widehat{F}^{-1}(p) = y_i \quad \text{where} \quad i-1 < pn \leq i. \quad (15)$$

For all quantiles $(j-1)/n < p \leq j/n$, the estimator of (2) is given by:

$$\widehat{\Lambda}^k(p) := \widehat{\Lambda}^k\left(\frac{j}{n}\right) = \frac{1}{(k-2)!}\frac{1}{n}\sum_{i=1}^{j}\left(\frac{j-i}{n}\right)^{k-2} y_i, \quad \forall k \geq 2. \quad (16)$$

The conditional S-Gini SEF can be calculated using the following approximation, resulting from the assumption that sample realizations are distinct. Let $\widehat{p}_i = \widehat{F}(y_i)$ be the population share associated to observed value y_i, corresponding to $\widehat{p}_i = i/n$ in this case. Using the fact that $\int_0^1 = \sum_{j=1}^n \int_{p_{i-1}}^{p_i}$ with $p_0 = 0$, the sample estimator of the conditional

S-Gini SEF can be written as follows:

$$\begin{aligned}
\widehat{W}^k(q,\widehat{F}) &= k\int_0^1 (q-p)^{k-1}\widehat{F}^{-1}(p)\mathbf{1}(p\leq q) \\
&= \sum_{i=1}^n y_i \int_{\widehat{p}_{i-1}}^{\widehat{p}_i} k(q-p)^{k-1}\mathbf{1}(p\leq q)dp \\
&= \sum_{i=1}^n y_j \left\{(q-\widehat{p}_{i-1})^k-(q-\widehat{p}_i)^k\right\}\mathbf{1}(i\leq qn)+y_{[qn]}(q-\widehat{p}_{[qn]})\mathbf{1}(qn\neq[qn]) \\
&\approx \sum_{i=1}^n y_i \left\{\left(\frac{n-i+1}{n}\right)^k-\left(\frac{n-i}{n}\right)^k\right\}\mathbf{1}(i\leq [qn]). \qquad (17)
\end{aligned}$$

Unless $p_i = q$, the last line is only an approximation of $\widehat{W}^k(q,\widehat{F})$, since there remain a residual proportional to $[qn]-i$ which vanishes as soon as the sample grows. Given the large size of the sample used in our applications, we take the empirical estimators (17) as the consistent estimator of $W^k(q,F)$.

The consistent estimator of the asymptotic covariance matrix of the conditional S-Gini SEF, $\mathbf{\Sigma}^k/n = \frac{1}{n}E\left[\boldsymbol{\phi}_i\cdot\boldsymbol{\phi}_i^t\right]$ can be obtained by replacing the population influence functions with the sample influence functions estimators. In the case of simple random sampling schemes, the empirical counterpart becomes:

$$\frac{\widehat{\mathbf{\Sigma}}^k}{n} = \frac{\frac{1}{n}\sum_{i=1}^n \boldsymbol{\phi}_i\cdot\boldsymbol{\phi}_i^t}{n}.$$

With a more complicated sampling rule (as in the case of stratified or clustered samples) the estimator must be computed using a non-uniform weighting schemes, as discussed by Zheng (2002). From Proposition 2, the estimator of the covariances between $\widehat{W}^k(q_j,\widehat{F})$ and $\widehat{W}^k(q_{j'},\widehat{F})$ coincides with the covariances of the sums of the influence functions of $\widehat{W}^k(q,\widehat{F})$ associated to q_j and $q_{j'}$, which gives (since the influence function average is zero):

$$\frac{\widehat{\varsigma}^k(q_j,q_{j'})}{n} = \frac{\frac{1}{n}\sum_{i=1}^n \phi_i\left(\widehat{W}^k(q_j,\widehat{F})\right)\phi_i\left(\widehat{W}^k(q_{j'},\widehat{F})\right)}{n}.$$

To compute the empirical covariances it is necessary to estimate the influence functions $\phi_i\left(\widehat{W}^k(q_j,\widehat{F})\right)$ of $\widehat{W}^k(q_j,\widehat{F})$ for every q_j, defined in (8). I make use of the results in Barrett and Donald (2009) to derive the conditional S-Gini SEF influence functions estimators, denoted $\widehat{\phi}_i\left(\widehat{W}^k(q,\widehat{F})\right)$. Following Barrett and Donald, the estimator $\widehat{\phi}_i(q,\widehat{GL})$ of the GL curve influence functions is:

$$\widehat{\phi}_i(q,\widehat{GL}) = (q\widehat{F}^{-1}(q)-\widehat{GL}(q))-\mathbf{1}(y_i\leq\widehat{F}^{-1}(q))(\widehat{F}^{-1}(q)-y_i). \qquad (18)$$

Substituting (18) into (8) gives the empirical counterpart of the influence function of

$\widehat{W}^k(q, \widehat{F})$, which can be decomposed into three components:

$$\widehat{\phi}_i\left(\widehat{W}^k(q, \widehat{F})\right) = \sum_{h=1}^{3} I_h.$$

Depending on the value of q and on the sample size, it is possible to derive either exact or approximate (by a number equal to $qn - [qn]$) empirical estimators. Since the correction term vanishes when the sample size is large, we only provide the formulation of the approximate estimators. These terms are:

i) $I_1 := \sum_{j=1}^{n} y_j \left[-k\left(\frac{j}{n}\left(\frac{[qn]-j}{n}\right)^{k-1} - \frac{j-1}{n}\left(\frac{[qn]-j+1}{n}\right)^{k-1}\right) - d_j(k)\right] \mathbf{1}(j \leq [qn])$,

ii) $I_2 := \sum_{j=1}^{n} \left[k\left(\left(\frac{[qn]-j}{n}\right)^{k-1} \widehat{GL}(j/n) - \left(\frac{[qn]-j+1}{n}\right)^{k-1} \widehat{GL}((j-1)/n)\right) + d_j(k)\right] \mathbf{1}(j \leq [qn])$,

iii) $I_3 := \sum_{j=1}^{n} -k\,(y_i - y_j)\,\mathbf{1}(y_i \leq y_j)\,[d_j(k-1)]\,\mathbf{1}(j \leq [qn])$,

where $d_j(\alpha) = \left(\frac{[qn]-j}{n}\right)^{\alpha} - \left(\frac{[qn]-j+1}{n}\right)^{\alpha}$ with α a positive natural number and $\widehat{GL} = \widehat{\Lambda}^2$ defined in (16).

6 Null hypothesis and test statistics for ISD and equality

This section makes use of the convergence results in (6) and (12) to derive tests statistics (and their limiting distributions) that permit to establish if distribution F ISD at order k the distribution G. Similarly to many other contributions dealing with estimation of stochastic orders (see for instance Barrett and Donald 2003), I assume ISDk ($F \succcurlyeq_{ISDk} G$) as the null hypothesis. According to the direct estimator approach, ISDk imposes restrictions on integrals of the quantile functions at every quantile p. Hence for $k \geq 1$:

$$\begin{array}{lll} H_0^k: & \Lambda_F^k(p) \geq \Lambda_G^k(p) & \text{for all } p \in [0,1]; \\ H_1^k: & \Lambda_F^k(p) < \Lambda_G^k(p) & \text{for some } p \in [0,1]. \end{array}$$

One can easily test for equality by reversing the role of F and G and testing if dominance is accepted also in this case.

According to the threshold estimator, ISDk imposes restrictions on the conditional S-Gini SEF at every quantile q. The null and alternative hypothesis for ISD at order $k \geq 3$ can be stated as follows:

$$\begin{array}{lll} H_0^k(W): & W^{k-1}(q,F) \geq W^{k-1}(q,G) & \text{for all } q \in [0,1]; \\ H_1^k(W): & W^{k-1}(q,F) < W^{k-1}(q,G) & \text{for some } q \in [0,1]. \end{array}$$

In practice, however, dominance can be tested only for a finite number m of percentiles. I take a similar stance as in Dardanoni and Forcina (1998, 1999) and Lefranc et al. (2009), among others, by constructing direct tests of dominance for a finite number m of linear constraints. Therefore the null assumption $F \succcurlyeq_{ISDk} G$ can be expressed as $\mathbf{\Lambda}_F^k - \mathbf{\Lambda}_G^k \geq 0$, which leads to the direct estimator approach for ISDk. Alternatively, the null of dominance can be expressed as $\mathbf{W}_F^{k-1} - \mathbf{W}_G^{k-1} \geq 0$, which leads to the threshold estimator approach developed in this paper.

6.1 Setting

To simplify the exposition, $D = F, G$ denotes a distribution while $\widehat{\mathbf{\Theta}}_D^k$ identifies the sample estimator used to test ISDk. The direct estimator approach requires to set $\widehat{\mathbf{\Theta}}_D^k = \widehat{\mathbf{\Lambda}}_D^k$ for $k \geq 1$, while the threshold estimator approach sets $\widehat{\mathbf{\Theta}}_D^k = \widehat{\mathbf{W}}_D^{k-1}$ for $k \geq 3$.

Let $\mathbf{\Theta}^k$ be the $2m \times 1$ vector obtained by staking the vectors $\mathbf{\Theta}_F^k$ and $\mathbf{\Theta}_G^k$. The sample counterpart of $\mathbf{\Theta}^k$ is denoted $\widehat{\mathbf{\Theta}}^k$ and is obtained from sub-samples of size n_F and n_G respectively, where $n = n_F + n_G$ indicates the pooled sample size. The scalars $r_F = n_F/n$ and $r_G = n_G/n$ indicate the relative size of each of the two samples compared to the pooled sample.

The hypothesis of dominance can be reformulated as a sequence of m linear constraints placed on the vector $\mathbf{\Theta}^k$. Let $\mathbf{R} = (\mathbf{I}_m, -\mathbf{I}_m)$ be the $m \times 2m$ *differences* matrix, with $\mathbf{I}_m$ indicating the $m \times m$ identity matrix. Define the parametric vector $\boldsymbol{\delta}_k \in \mathbb{R}^m$ as:

$$\boldsymbol{\delta}_k \;=\; \mathbf{R}\mathbf{\Theta}^k.$$

I maintain the (non testable) assumption that F and G are generated by independent processes. The various hypothesis of dominance or equality can be written in terms of linear inequalities involving $\boldsymbol{\delta}_k$, with $\widehat{\boldsymbol{\delta}}_k$ indicating its empirical counterpart. Using the results in (6) or (12) and under the independence assumption, the following asymptotic result holds for every $k \geq 1$:

$$\sqrt{n}\,\widehat{\boldsymbol{\delta}}_k \;=\; \sqrt{n}\,\mathbf{R}\,\widehat{\mathbf{\Theta}}^k \;\text{ is asymptotically distributed as }\; \mathcal{N}\left(\sqrt{n}\,\mathbf{R}\,\mathbf{\Theta}^k,\; \mathbf{\Omega}\right), \tag{19}$$

where $\widehat{\boldsymbol{\delta}}_k$ denotes the sample estimator of $\boldsymbol{\delta}_k$, and

$$\mathbf{\Omega} \;=\; \mathbf{R}\,\mathrm{diag}\left(\frac{\mathbf{\Sigma}_F^k}{r_F}, \frac{\mathbf{\Sigma}_G^k}{r_G}\right)\mathbf{R}^t.$$

According to the direct estimator approach, ISDk empirical implementation is possible by using the estimator of the asymptotic covariance of $\widehat{\mathbf{\Lambda}}_D^k$ as an estimator for the asymptotic variance $\widehat{\mathbf{\Omega}}$. According to the threshold estimator approach, $\widehat{\mathbf{\Omega}}$ coincides with the estimator of the asymptotic covariance of $\widehat{\mathbf{W}}_D^{k-1}$ in (10). The asymptotic normality of the estimators permits to test dominance through Wald-type tests statistics.

6.2 Testing equality

In the case of equality testing, the null and alternative hypothesis are:

$$H_0^k: \ \boldsymbol{\delta}_k = \mathbf{0} \qquad H_1^k: \ \boldsymbol{\delta}_k \neq \mathbf{0}.$$

Under the null hypothesis, it is possible to resort to a Wald test static T_1^k:

$$T_1^k \ := \ n \, \widehat{\boldsymbol{\delta}}_k^t \, \widehat{\boldsymbol{\Omega}}^{-1} \, \widehat{\boldsymbol{\delta}}_k.$$

Given the convergence results in (19), the asymptotic distribution of the test T_1^k is χ_m^2. The p-value tabulation follows the usual rules.

6.3 Testing dominance

In the case of strong dominance testing, i.e. $F \succcurlyeq_{ISDk} G$, the null and alternative hypothesis for direct and threshold ISDk testing can be stated as follows:

$$H_0^k: \ \boldsymbol{\delta}_k \in \mathbb{R}_+^m \qquad H_1^k: \ \boldsymbol{\delta}_k \not\in \mathbb{R}_+^m.$$

The Wald-type test statistic under inequality constraints has been developed by Kodde and Palm (1986). Under the null H_0^k the test statistics T_2^k is:

$$T_2^k = \min_{\boldsymbol{\delta}_k \in \mathbb{R}_+^m} \left\{ n \, (\widehat{\boldsymbol{\delta}}_k - \boldsymbol{\delta}_k)^t \, \widehat{\boldsymbol{\Omega}}^{-1} \, (\widehat{\boldsymbol{\delta}}_k - \boldsymbol{\delta}_k) \right\}.$$

Kodde and Palm (1986) have shown that the statistic T_2^k is asymptotically distributed as a mixture of χ^2 distributions, provided that the asymptotic normality in (19) holds:

$$T_2^k \ \sim \ \overline{\chi}^2 \ = \ \sum_{j=0}^{m} w\left(m, m-j, \widehat{\boldsymbol{\Omega}}\right) \Pr\left(\chi_j^2 \geq c\right),$$

where $w\left(m, m-j, \widehat{\boldsymbol{\Omega}}\right)$ denotes the probability that $m-j$ elements of $\boldsymbol{\delta}_k$ are strictly positive.[4]

To test the reverse dominance order, that is $G \succcurlyeq_{ISDk} F$, it is sufficient to replace $-\widehat{\boldsymbol{\delta}}_k$ and $-\boldsymbol{\delta}_k$ in the calculation of T_2^k.

7 Illustration: Equality of opportunity in France

This section provides an illustrative application of ISD to check for robustness in *equality of opportunity* (EOP hereafter) assessments. EOP for income acquisition posits that

[4] To estimate $w\left(m, m-j, \widehat{\boldsymbol{\Omega}}\right)$, we draw 10,000 multivariate normal vectors and covariance matrix $\widehat{\boldsymbol{\Omega}}$, provided it is positive definite. Then we compute the proportion of vectors with $m-j$ positive elements.

differences in family background across individuals should not predict their labor market income prospects (Fleurbaey 2008).

Let assume for simplicity that there are only two distinct family backgrounds a and b, addressed to as individual circumstances $c \in \{a, b\}$. Within this setting, the income y_i of individual i is determined by two components, her circumstances c and her effort e, so that $y_i = y(c, e)$. If the two components are independent, the effort exerted by i can be associated to her position in the ranking of incomes, let say p, so that $p = e$ (as in Roemer 1998) and the individual income can be estimated by the overall income distribution. Let $\widehat{F}$ be the empirical income process, it follows that $y_i = \widehat{F}^{-1}(p|c)$.

According to the Lefranc et al. (2009) view, EOP prevails if there is no agreement among inequality averse social evaluation functions in preferring a society where incomes are distributed according to $\widehat{F}^{-1}(p|a)$ rather than $\widehat{F}^{-1}(p|b)$ or *vice-versa.* That is:

$$\widehat{F}(.|c) \not\succcurlyeq_{ISD2} \widehat{F}(.|c') \text{ with } c \neq c'.$$

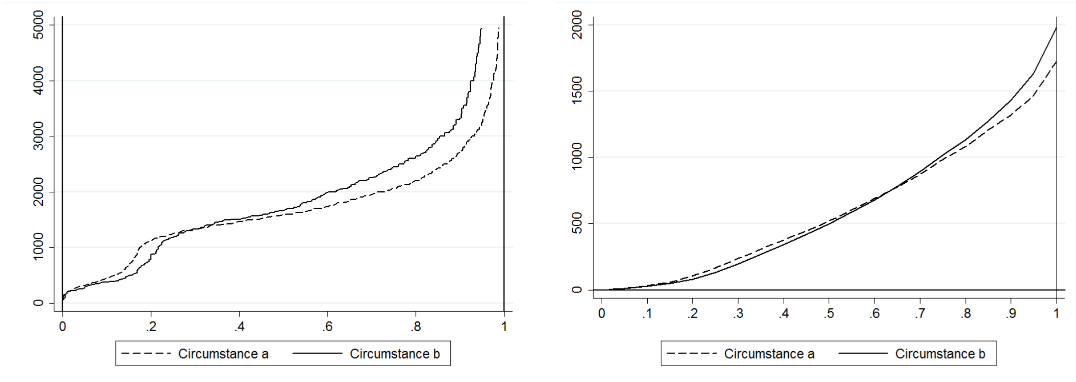
When EOP prevails, the cases $\widehat{F}(.|c) \succcurlyeq_{ISD3} \widehat{F}(.|c')$ and $\widehat{F}(.|c) \succcurlyeq_{ISDk} \widehat{F}(.|c')$ with k arbitrary large are equivalent. However, there is larger agreement on the existence of an unjust advantage of c over c' whenever ISD3 holds than there is whenever ISDk holds. Hence, if EOP is accepted it is essential to test at least for ISD3 in order to understand how much robust is the violation of ISD2.

To qualify the degree of EOP robustness, I make use of ISD tests under the hypothesis of equality or dominance at order k. Practically, a test for EOP robustness requires (i) to partition the sample into groups defined by circumstances a and b, (ii) to determine the groups specific distributions $\widehat{F}(.|a)$ and $\widehat{F}(.|b)$ and (iii) to check the minimal order k at which $\widehat{F}(.|a) \succcurlyeq_{ISDk} \widehat{F}(.|b)$ or $\widehat{F}(.|b) \succcurlyeq_{ISDk} \widehat{F}(.|a)$ or both. The focus will be exclusively on ISD3 comparisons.

I make use of the FLFS - French Labor Force Survey data (*Enquête Emploi*) provided by INSEE to estimate the labor income prospects made conditional on family background characteristics: the *circumstance a* gathers all French workers whose parents are either non-french or were occupied as manual workers or farmers; the *circumstance b* is instead associated to the middle class parental background, gathering artisans, small entrepreneurs and non-manual workers.[5] The investigations are restricted to FLFS waves 2004, 2006, 2008 and 2010 and to cohorts born between 1958 and 1962 to approximate the expected earning profile with an homogeneous sample in terms of age patterns and experience in the labor market.[6] Note that the approach is fully distributional and it does not involve the comparisons of individual realizations.

[5]The higher social class, comprising white collars, managers, professionals is not considered in this illustrative example.

[6]The FLFS is a rotating panel. The panel rotation after 2003 is of one year and a half (that is, one-sixth of the sample is replaced every trimester). Picking up information every two years allows to deal with a renewed sample, as in years 2004, 2006, 2008 and 2010. The year 2002 is not exploited due to imperfections in the data collected.



(a) Pen's Parade (b) Generalized Lorenz curves

Figure 2: ISD1 and ISD2 relations between circumstance a an circumstance b.

The estimating sample consists of $2,326$ French workers ($1,810$ observations are associated to circumstance a and 516 to circumstance b). Figure 2 reports the Pen's Parade and the generalized Lorenz curves of the two circumstances' incomes distributions.[7] The two income profiles cannot be ranked according to ISD1 or ISD2 since their respective quantile functions and the generalized Lorenz curves cross at least once. It seems clear, however, that the poorest workers in group a enjoy an higher advantage than the poorest workers in group b, while the order swaps as soon as the income deciles grow. This relation suggests that the correct dominance relation to verify is $\widehat{F}(.|a) \succcurlyeq_{ISDk} \widehat{F}(.|b)$ for $k = 1, 2, 3$.

The FLFS data are *areolar*: they are not drawn directly from a selection of households or individuals, but from a selection of geographical areas made up of twenty adjacent households on average. Then, information on earnings for workers aged 15 to 65 within each area is collected in the survey. I take the clustered sampling scheme of the FLFS into account when computing the covariance structure of the influence functions associated to the conditions S-Gini social evaluation functions calculated at different income deciles.

Table 1 reports incomes quantiles, generalized Lorenz curves coordinates, Λ^3 coordinates as well as conditional S-Gini SEF $W^3(., F)$ indices for selected deciles of $F(.|a)$ and $F(.|b)$, calculated accounting for the survey design of FLFS only for the threshold estimator approach. The table shows that, independently on the evaluation method used to infer ISD3, pairwise comparisons of estimated coefficients at fixed deciles leads to inconclusive results: in many cases the differences in quantiles curves, generalized Lorenz curves, integrals of the generalized Lorenz curves or conditional S-Gini SEF at a given decile are statistically zero, or their signs do not concord across deciles. Joint tests for the equality and dominance null hypothesis are therefore preferred.

The Wald-type test statistics and their simulated p-values are reported in Table 2. Joint tests show that rank and generalized Lorenz equality/dominance nulls are rejected at 1%. Thus EOP is satisfied and ISD comparisons at higher orders become relevant. The

[7]Data have been depurated of years of survey fixed effects.

Table 1: Statistics and standard errors for $\widehat{F}(.|a) \succcurlyeq_{ISDk} \widehat{F}(.|b)$.

	Wage, circumstance a (1)		Wage, circumstance b (2)		Difference (1)-(2)
	Coeff.	*(SE)*	*Coeff.*	*(SE)*	
Rank dominance					
10%	446.3	(963.51)	378.8	(1188.67)	67.44
30%	1327.3	(736.97)	1327.3	(1200.39)	0.00
50%	1593.9	(707.16)	1666.5	(1107.57)	-72.63
70%	1950.4	(1009.56)	2253.0	(1529.40)	-302.66*
90%	2700.0	(2258.25)	3319.7	(3418.10)	-619.74*
Generalized Lorenz dominance					
10%	31.5	(47.68)	28.0	(39.29)	3.41
30%	230.1	(337.98)	196.5	(375.59)	33.62
50%	522.0	(429.43)	500.1	(498.45)	21.86
70%	871.9	(535.44)	893.6	(690.61)	-21.71
90%	1321.1	(695.79)	1430.8	(935.84)	-109.71*
ISD at order 3 (threshold estimator)					
10%	2.7	(4.79)	2.2	(4.08)	0.41
30%	48.2	(104.32)	38.6	(80.26)	9.63*
50%	196.8	(291.01)	175.9	(264.26)	20.85
70%	472.8	(546.76)	448.9	(513.15)	23.97
90%	906.3	(896.97)	904.0	(854.19)	2.38
ISD at order 3 (direct estimator)					
10%	1.3	(3.40)	1.2	(2.97)	0.17
30%	24.1	(65.46)	19.5	(54.17)	4.60
50%	98.4	(165.54)	88.0	(167.98)	10.38
70%	236.4	(308.29)	225.4	(340.01)	11.05
90%	453.2	(527.69)	453.5	(608.52)	-0.28
ISD at order 3 (bootstrap estimator)					
10%	2.7	(0.25)	2.3	(0.35)	0.40*
30%	48.2	(2.09)	38.6	(3.04)	9.58*
50%	196.8	(6.35)	176.0	(9.78)	20.76*
70%	472.9	(11.47)	449.0	(16.83)	23.84*
90%	906.4	(17.09)	904.2	(24.52)	2.21
Mean	1727.0	(31.72)	1984.9	(70.05)	-257.86*
Sample size:	1810		516		-

Notes: FLFS data, waves 2004, 2006, 2008 and 2010. The *threshold* and *bootstrap* estimators report conditional S-Gini coordinates for the selected deciles. Standard errors reported in parenthesis account for the clustering structure of the sample (*aires*). Bootstrapped standard errors with 200 repetitions.
* denotes significantly different from zero at 5% on the basis of a two-tailed test.

threshold estimator, the preferred inference strategy, shows that ISD3 cannot be rejected by the data at 1% significance level, while equality of conditional S-Gini SEF vectors is clearly rejected at any significance level. Interesting, the bootstrapping procedure (which is computationally less intensive) leads to similar conclusions as the threshold estimator. Both procedures permit to take into account the clustering design the data. The direct estimator for ISD3, instead, consider data as obtained from a simple random sampling scheme. As a consequence the direct estimator rejects dominance, while equality of the integrals of the generalized Lorenz curves cannot be rejected even at 10% significance level.

Table 2: Wald tests and p-values for $\widehat{F}(.|a) \succcurlyeq_{ISDk} \widehat{F}(.|b)$.

Null hypothesis:	Equality (1)		Dominance (2)	
	Wald test	*p-value*	*Wald test*	*p-value*
Rank	55.13	0.000	41.35	0.000
Generalized Lorenz	45.33	0.000	17.91	0.000
ISD order 3				
- Threshold est.	842.75	0.000	0.00	0.652
- Direct est.	4.55	0.871	0.00	1.000
- Bootstrap est.	8600.90	0.000	0.00	0.951

Notes: FLFS data, waves 2004, 2006, 2008 and 2010. Wald-type tests for equality and dominance based on distributions deciles. The *threshold* and *bootstrap* tests are performed over conditional S-Gini coordinates. P-values for equality obtained from a Chi-squared distribution. P-values for the dominance hypothesis are obtained via Monte-Carlo simulations. The null hypothesis of dominance cannot be rejected (is rejected) at the significance level α if the Wald test is lower (larger) than the corresponding *lower bound* lb_α (*upper bound* ub_α). For $\alpha = 10\%,\ 5\%, 1\%$ the corresponding lower bounds are $lb_\alpha = 1.642,\ 2.706,\ 5.412$ and the corresponding upper bounds are $ub_\alpha = 14.067,\ 16.274,\ 20.972$ (see Kodde and Palm 1986).

To conclude, EOP cannot be rejected. However, EOP seems not to be robust since the two distributions associated to circumstances a and b can be ranked according to ISD3.

8 Concluding remarks

Inverse Stochastic Dominance (ISD) is a convenient tool for constructing robust welfare comparisons between pairs of distributions. Using the parallel with the dual theory of choice under risk (Yaari 1987), income distributions play the role of risky lotteries. ISD impose empirically testable criteria on the data that permit to identify the class of social evaluation functions where agreement is reached over the preferred income lottery. One important class of social evaluation functions originates from the Single Parameter Gini (S-Gini) index family (Donaldson and Weymark 1983), parametrized by the ISD order of dominance. For instance, the Gini index can be related to ISD at order three. I propose an innovative procedure for constructing ISD dominance at orders higher than the second by setting dominance conditions among conditional single parameters Gini social evaluation functions, calculated at various quantiles of the distribution. Standard errors are computed by calculating the influence functions of the conditional S-Ginis SEF for each distributions. Differently from other approaches, the influence functions permit to account for non-standard randomization procedures in sample data collection, a characteristic common to many economic surveys. One of these is the French Labor Force Survey, which is used to test the robustness of the Equality of Opportunity evaluations. The ISD tests proposed in this paper are operationalized through the user written `isdtest` Stata routine.

References

Aaberge, R. (2006). Asymptotic distribution theory of empirical rank-dependent measures of inequality, *in* V. Nair (ed.), *Advances in Statistical Modeling and Inference – Essays in Honor of Kjell A. Doksum*, World Scientific, chapter 5.

Aaberge, R. (2007). Gini's nuclear family, *The Journal of Economic Inequality* **5**(3): 305–322.

Aaberge, R. (2009). Ranking intersecting lorenz curves, *Social Choice and Welfare* **33**: 235–259.

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

Barrett, G. F. and Donald, S. G. (2009). Statistical inference with generalized Gini indices of inequality, poverty, and welfare, *Journal of Business & Economic Statistics* **27**(1): 1–17.

Beach, C. M. and Davidson, R. (1983). Distribution-free statistical inference with Lorenz curves and income shares, *The Review of Economic Studies* **50**(4): 723–735.

Cowell, F. (2000). Chapter 2 measurement of inequality, Vol. 1 of *Handbook of Income Distribution*, Elsevier, pp. 87 – 166.

Dardanoni, V. and Forcina, A. (1998). A unified approach to likelihood inference on stochastic orderings in a nonparametric context, *Journal of the American Statistical Association* **93**(443): 1112–1123.

Dardanoni, V. and Forcina, A. (1999). Inference for Lorenz curve orderings, *Econometrics Journal* **2**: 49–75.

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.

Davies, J. and Hoy, M. (1995). Making inequality comparisons when lorenz curves intersect, *The American Economic Review* **85**(4): pp. 980–986.

Donaldson, D. and Weymark, J. A. (1983). Ethically flexible Gini indices for income distributions in the continuum, *Journal of Economic Theory* **29**(2): 353 – 358.

Fishburn, P. C. (1976). Continua of stochastic dominance relations for bounded probability distributions, *Journal of Mathematical Economics* **3**(3): 295 – 311.

Fishburn, P. and Vickson, R. G. (1978). Theoretical foundations of stochastic dominance, *in* G. Whitmore and M. Findlay (eds), *Stochastic Dominance: and Approach to Decision-Maing Under Risk*, D.C. Heat, Lexington, chapter 2, pp. 39–114.

Fleurbaey, M. (2008). *Fairness, Responsability and Welfare*, Orxford University Press.

Foster, J. E. and Shorrocks, A. F. (1988). Poverty orderings, *Econometrica* **56**(1): pp. 173–177.

Foster, J., Greer, J. and Thorbecke, E. (1984). A class of decomposable poverty measures, *Econometrica* **52**(3): 761–66.

Gastwirth, J. L. (1971). A general definition of the Lorenz curve, *Econometrica* **39**(6): 1037–1039.

Kodde, D. A. and Palm, F. C. (1986). Wald criteria for jointly testing equality and inequality restrictions, *Econometrica* **54**(5): pp. 1243–1248.

Kolm, S.-C. (1969). *The Optimal Production of Social Justice*, Public Economics, London: McMillan, pp. 145–200.

Kovacevic, M. S. and Binder, D. A. (1997). Variance estimation for measures of income inequality and polarization: The estimating equations approach, *Journal of Official Statistics* **13**(1): 41–58.

Lefranc, A., Pistolesi, N. and Trannoy, A. (2009). Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France, *Journal of Public Economics* **93**(11-12): 1189 – 1207.

Maccheroni, F., Muliere, P. and Zoli, C. (2005). Inverse stochastic orders and generalized Gini functionals, *METRON* **63**(3): 529–559.

Muliere, P. and Scarsini, M. (1989). A note on stochastic dominance and inequality measures, *Journal of Economic Theory* **49**(2): 314 – 323.

Roemer, J. (1998). *Equality of Opportunity*, Harvard University Press, Cambridge.

Saposnik, R. (1981). Rank-dominance in income distributions, *Public Choice* **36**(1): 147–151.

Shorrocks, A. F. (1983). Ranking income distributions, *Economica* **50**(197): pp. 3–17.

Yaari, M. E. (1987). The dual theory of choice under risk, *Econometrica* **55**(1): pp. 95–115.

Zheng, B. (1999). Statistical inferences for testing marginal rank and (generalized) Lorenz dominances, *Southern Economic Journal* **65**(3): 557–570.

Zheng, B. (2002). Testing Lorenz curves with non-simple random samples, *Econometrica* **70**(3): 1235–1243.

Zoli, C. (1999). Intersecting generalized Lorenz curves and the Gini index, *Social Choice and Welfare* **16**: 183–196.

Zoli, C. (2002). Inverse stochastic dominance, inequality measurement and Gini indices, *Journal of Economics* **9**: 119–161.