DISCUSSION PAPER SERIES

IZA DP No. 7738

A Theory for Ranking Distribution Functions

Rolf Aaberge
Tarjei Havnes
Magne Mogstad

November 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

A Theory for Ranking Distribution Functions

Rolf Aaberge
Statistics Norway,
ESOP and IZA

Tarjei Havnes
University of Oslo,
ESOP, Statistics Norway and IZA

Magne Mogstad
University College London,
Statistics Norway, ESOP and IZA

Discussion Paper No. 7738
November 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7738
November 2013

ABSTRACT

A Theory for Ranking Distribution Functions

When is one distribution (of income, consumption, or some other economic variable) more equal or better than another? This question has proven difficult to answer in situations where distribution functions intersect and no unambiguous ranking can be attained without introducing weaker criteria than second-degree stochastic dominance. The conventional approach in empirical work is to adopt some summary statistics, with no explicit reason being given for preferring one measure rather than another. In this paper, we develop a theory for ranking distribution functions. Our theory offers a general framework to unambiguously rank any set of distribution functions and quantify the social welfare level of a dominating distribution as compared to a dominated distribution. The framework is based on two complementary sequences of nested dominance criteria. The first (second) sequence extends second-degree stochastic dominance by placing more emphasis on differences that occur in the lower (upper) part of the distribution. These sequences of dominance criteria characterize two separate systems of nested subfamilies of social welfare functions. This allows us to identify the least restrictive social preferences that give an unambiguous ranking of any set of distribution functions. We also provide an axiomatization of the sequences of dominance criteria and the corresponding subfamilies of social welfare functions. To perform inference, we develop asymptotic distribution theory for empirical dominance criteria where it is demonstrated that the associated empirical processes converge in distribution to Gaussian processes. The usefulness of our framework is illustrated with two empirical applications; the first assesses the social welfare implications of changes in household income distributions over the business cycle, while the second ranks the actual and counterfactual outcome distributions from a policy experiment.

JEL Classification: D30, D63, I31

Keywords: distribution functions, stochastic dominance, social welfare, inequality

Corresponding author:

Rolf Aaberge
Research Department
P.O. Box 8131 Dep.
Statistics Norway
0033 Oslo
Norway
E-mail: rolf.aaberge@ssb.no

1 Introduction

How do we compare intersecting distribution functions? The answer to this question is important for both descriptive analysis and policy evaluation. A key task of statistical offices and government agencies is to compare distribution functions of economic variables across countries, subgroups and time. Much descriptive research is about analyzing changes in and differences between distributions of wages, income, consumption and wealth, as they are considered important determinants of economic welfare as well as markers for what kind of activities are rewarded in an economy. There is also a growing body of research on how to assess the distributional effects of policy changes: The literature has developed methods for estimating the counterfactual outcome distribution in the absence of a policy intervention,[1] but has generally stopped short of establishing a framework for ranking the actual and counterfactual outcome distributions.

In this paper, we develop a theory for ranking distribution functions. Our theory offers a general framework to unambiguously rank any set of distribution functions and quantify the social welfare level of a dominating distribution as compared to a dominated distribution. Since the seminal contributions of Kolm (1969) and Atkinson (1970), second-degree stochastic dominance has become a widely accepted criterion for ranking distribution functions. But in many applications where the distribution functions intersect, a reasonable refinement of this criterion is necessary to attain an unambiguous ranking.[2] Although the theoretical literature offers dominance criteria of third or higher degree,[3] they are rarely used; the reason is that higher degree dominance criteria are often viewed as difficult to interpret and hard to justify because they rely on assumptions about third or higher order derivatives (see e.g. Atkinson, 2003, 2008). Thus, most empirical studies consider a few moments or use a parametric social welfare function when ranking intersecting distribution functions. A natural concern is that the conclusions reached in these studies are sensitive to the choice of moments or specification of social welfare function.[4]

Our framework for comparing intersecting distribution functions is based on two complementary sequences of nested inverse stochastic dominance criteria.[5] The first sequence includes

[1]For example, a number of papers have focused on identification and estimation of unconditional quantile treatment effects under unconfoundedness (e.g. Firpo, 2007; Firpo, Fortin, and Lemieux, 2009) or with selection on unobservables (e.g. Imbens and Newey, 2009; **?**). See e.g. Bitler, Gelbach, and Hoynes (2006, 2008) for empirical evaluations of the distributional effects of policy interventions.

[2]Several studies have demonstrated the limited practical scope for ranking income distributions according to second-degree stochastic dominance (see e.g. Davies and Hoy, 1995; Atkinson, 2008).

[3]See e.g. Fishburn (1976), Fishburn (1980), Chew (1983), and Fishburn and Willig (1984) for extensions of stochastic dominance to an arbitrary order.

[4]The challenge in ranking distribution functions by their moments is twofold. First, the moments of an unbounded distribution do not uniquely determine the distribution function. For example, there exists several distributions with the same moments as the log-normal distribution (**?**). Second, it is not clear how to aggregate and weigh the various moments of the distributions being compared.

[5]While second-degree inverse stochastic dominance is equivalent to second-degree stochastic dominance (Hardy, Littlewood, and Pólya, 1934; Kolm, 1969; Atkinson, 1970), the two types of dominance differ at the third or higher degree. See e.g. Le Breton and Peluso (2009) for a discussion.

the traditional inverse dominance criteria of third and higher degrees; it is called *upward dominance* because it aggregates the inverse of the distribution function from below, and therefore places more emphasis on differences that occur in the lower part of the distribution. The second sequence is novel and complements the traditional criteria by placing more emphasis on differences that occur in the upper part of the distribution; we call it *downward dominance* because it aggregates the integrated inverse distribution function from above. Since the sequences are hierarchical, the sensitivity to differences in the lower (upper) part of the distribution increases with the degree of upward (downward) dominance. The two sequences coincide at second-degree dominance, and thus both satisfy the Pigou-Dalton transfer principle.

For each sequence, we show that dominance of any degree can be given a simple social welfare interpretation. For example, ranking distribution functions according to third-degree upward dominance is equivalent to employing the Gini social welfare function to compare the welfare of individuals located in the lower tail of each quantile of the distributions.[6] As a consequence, we do not have to rely on assumptions about third and higher order derivatives to interpret the sequences of dominance criteria. To make statistical inference about upward and downward dominance of any degree, we develop asymptotic distribution theory for empirical dominance criteria where it is demonstrated that the associated empirical processes converge in distribution to Gaussian processes. Thus, the empirical dominance criteria are asymptotically normally distributed both when considered as processes and for fixed ranks in the distribution.[7]

We next characterize the relation between upward and downward dominance and social welfare functions in the ranking of distribution functions. For each sequence, we show equivalence in the ranking of distributions according to the dominance criteria and a general family of rank-dependent social welfare functions. The family of rank-dependent social welfare functions was originally proposed by Yaari (1987; 1988), and can be represented as weighted averages of the outcomes of interest where the weight decreases with the rank in the outcome distribution. The functional form of the weighting function details the inequality aversion of a social planner who employs the family of social welfare functions to compare intersecting distribution functions. Because the sequences of dominance criteria are nested, our equivalence results allow us to uniquely identify the largest subfamily of welfare functions – and thus the least restrictive social preferences – that give an unambiguous ranking of any set of distribution functions.

We also provide a characterization of the largest subfamily of social welfare functions that rank consistently with dominance of any given degree. Because of the equivalence result, this characterization gives a normative justification not only for the social welfare functions, but also for the use of higher degree dominance criteria when comparing distribution functions. The subfamily associated with upward dominance is characterized by (generalizations of) the

[6] The Gini social welfare function was originally introduced by Sen (1974), and was given a complete axiomatic justification by Aaberge (2001).

[7] We are not aware of asymptotic distribution theory for inverse stochastic dominance tests. See, for example, Abadie (2002), Anderson (1996), Barrett and Donald (2003), Linton, Maasoumi, and Whang (2005), and Davidson and Duclos (2000) for alternative approaches to testing for standard stochastic dominance.

principle of downside positional transfer sensitivity (see Zoli, 1999; Aaberge, 2000; 2009), while the subfamily associated with downward dominance is characterized by (generalizations of) the principle of upside positional transfer sensitivity (see Aaberge, 2009). The two principles differ in the sensitivity to differences in the lower versus upper part of the distribution.

To not only answer whether one distribution is better than another distribution, but also get an estimate of by how much, it is convenient to work with parametric social welfare functions. We show that the members of two alternative parametric families of social welfare functions can be divided into subfamilies according to their relationship with the nested inverse stochastic dominance criteria. The parametric family that ranks consistently with upward (downward) dominance criteria exhibits successively higher aversion to differences in the lower (upper) part of the distribution. The parametric families are well known, easily implementable and the estimated social welfare can be given a money metric interpretation. Since each family uniquely determines the distribution function, no information is lost by restricting focus to these parametric social welfare functions.

We show the usefulness of our framework using two empirical applications. The first application uses data from the UK to study how the distribution of household income evolved over a boom and a bust era in the British economy. We show how our framework can be used to make unambiguous statements about the social welfare implications of the changes in the household income distribution over the business cycle. The second application uses random-assignment data to evaluate the distributional effects of Connecticut's Jobs First program, which involved generous earnings disregard and strict time limits.[8] We use our framework to infer the least restrictive social preferences that allow an unambiguous conclusion of whether this program was an overall success. In both applications, we find that third-degree downward dominance is a particularly powerful refinement of second-degree dominance, providing an almost complete ranking of the distribution functions. By comparison, the traditional criterion of third-degree upward dominance resolves few of the comparisons that were ambiguous under second-degree dominance.

Our paper is related to a growing literature on refinements of second-degree dominance in the comparison of distribution functions. In particular, much work has been done on third-degree dominance and its relationship to social welfare and inequality (for reviews, see Lambert, 1993; Le Breton and Peluso, 2009). One strand of the literature is influenced by expected utility theory and explores third-degree stochastic dominance as a criterion for ranking distributions. For example, Shorrocks and Foster (1987) consider third-degree stochastic dominance in the case of a single intersection of the Lorenz curves; Davies and Hoy (1995) study the general case of Lorenz curves with multiple intersections and show that for distributions with the same mean, third-degree stochastic dominance is equivalent to the comparison of variances for ap-

[8]Our choice to use the Jobs First program is not incidental: As shown in Bitler, Gelbach, and Hoynes (2006), the estimated quantile treatment effects exhibit the substantial heterogeneity predicted by labor supply theory. As a consequence, the distributions of income with and without the Jobs First program intersect.

propriate truncated income distributions.[9] Another strand of the literature exploits the ideas and techniques of non-expected utility theory to examine third-degree inverse stochastic dominance as a criterion for ranking distributions (see e.g. Muliere and Scarsini, 1989; Zoli, 1999; Zoli, 2002; Aaberge, 2009). Our paper contributes by exploring the relation between upward and downward inverse stochastic dominance of any degree and a general family of rank-dependent social welfare functions. Taken together, our results provide a general framework to unambiguously rank any set of distribution functions and to quantify the social welfare level of a dominating distribution as compared to a dominated distribution.

The remainder of the paper proceeds as follows. Section 2 characterizes the relationship between inverse stochastic dominance and social welfare functions as criteria for ranking distribution functions. Section 3 identifies and describes the parametric families that rank distributions consistent with upward and downward dominance. Section 4 presents the asymptotic distribution theory. Section 5 provides the empirical applications, before Section 6 concludes.

2 Inverse stochastic dominance and social welfare

This section begins by reviewing the relationship between second-degree dominance and the general family of social welfare functions. We next introduce upward and downward dominance of third degree as criteria for ranking distribution functions, and characterize their relationship to social welfare functions. Finally, we introduce the full hierarchical sequences of nested inverse stochastic dominance criteria, and show how they allow us to uniquely identify the largest subfamily of social welfare functions required to reach an unambiguous ranking of any set of distribution functions.

2.1 Second-degree dominance and rank-dependent welfare functions

Let F be a member of the set $\mathscr{F}$ of cumulative distribution functions with mean μ_F and left inverse defined by

$$F^{-1}(t) = \inf\{x : F(x) \geq t\}$$

Note that both discrete and continuous distribution functions are allowed in $\mathscr{F}$, and though the former is what we actually observe, the latter often allows simpler derivation of theoretical results and is a valid large sample approximation. Thus, in most cases below, F will be assumed to be a continuous distribution function, but the assumption of a discrete distribution function will be used where appropriate. To fix ideas, we will refer to F as the income distribution, although our framework can be applied to any type of distribution function.

[9]See also Shorrocks (1983), Atkinson (2008), Chiu (2007), Davies and Hoy (1994), Dardanoni and Lambert (1988), Le Breton and Peluso (2009), and Le Breton, Michelangeli, and Peluso (2012).

Second degree dominance

Since the seminal contributions of Kolm (1969) and Atkinson (1970), second-degree dominance has become a widely accepted criterion for ranking distribution functions.[10]

Definition 2.1. A distribution function F_1 is said to *second-degree dominate* a distribution function F_0 if and only if

$$\int_0^u F_1^{-1}(t)dt \geq \int_0^u F_0^{-1}(t)dt \text{ for all } u \in [0,1]$$

and the inequality holds strictly for some $u \in (0,1)$.

As is well known, all inequality averse social planners rank distribution functions consistently with second-degree dominance. But in many applications, weaker criteria than second-degree dominance are required to obtain an ordering of distributions.

Rank-dependent social welfare functions

As in the literature on choice under uncertainty, ranking criteria can be derived from independence axioms imposed on the ordering $\succeq$ defined on $\mathscr{F}$. The preference relation $\succeq$ of the social planner is assumed to be continuous, transitive and complete and to rank $F_1 \succeq F_0$ if $F_1^{-1}(t) \geq F_0^{-1}(t)$ for all $t \in [0,1]$. To give the preferences of the planner an empirical content, Yaari (1988; 1987) imposes the so-called dual independence axiom on $\succeq$, defined by

Axiom 1. ***(Dual Independence)***. *Let F_0, F_1 and F_2 be members of $\mathscr{F}$ and let $\alpha \in [0,1]$. Then $F_1 \succeq F_0$ implies $\left(\alpha F_1^{-1} + (1-\alpha)F_2^{-1}\right)^{-1} \succeq \left(\alpha F_0^{-1} + (1-\alpha)F_2^{-1}\right)^{-1}$.*

Armed with this axiom, Yaari (1987; 1988) proved that the preference relation $\succeq$ can be represented by the following rank-dependent family of social welfare functions

$$W_P(F) = \int_0^1 P'(t)F^{-1}(t)dt, \tag{2.1}$$

where P' is the derivative of a preference function from the following set.

$$\mathscr{P} = \left\{P : P'(t) > 0 \, and \, P''(t) < 0 \, for \, all \, t \in (0,1), P'(1) = P(0) = 0, P(1) = 1\right\}.$$

The dual independence axiom requires that the ordering $\succeq$ is invariant with respect to identical mixing of the *inverses of the distribution functions* being compared; that is, mixing of income levels given population shares. By comparison, the independence axiom used in Atkinson

[10] Since second-degree inverse stochastic dominance is equivalent to second-degree stochastic dominance, we will simply refer to this criterion as second-degree dominance.

(1970) requires that the ordering of *distribution functions* is invariant with respect to identical mixing of the *distributions* being compared; that is, mixing of population shares given income levels. For further discussion, see Yaari (1988) and Aaberge (2001).

Relation between second-degree dominance and rank-dependent welfare functions

As demonstrated by Yaari (1988), the social welfare functions W_P are consistent with the condition of second-degree stochastic dominance if and only if $P'(t) > 0$ and $P''(t) < 0$. It follows by straightforward calculations that $0 \leq W_P \leq \mu_F$ for strictly concave P and that $W_P = \mu_F$ if and only if F is the egalitarian distribution. Thus, W_P can be interpreted as the equally distributed equivalent income (see Atkinson, 1970). With equal means, the condition of second-degree stochastic dominance is identical to the Pigou-Dalton transfer principle, which states that an income transfer from a richer to a poorer individual reduces income inequality, provided that their ranks in the income distribution are unchanged.

The general family of social welfare functions W_P represents a preference relation defined on the set of distribution functions. The preference function P assigns weights to the incomes of the individuals in accordance with their rank in the income distribution. Therefore, the functional form of P reveals the attitude towards inequality of a social planner who employs W_P to judge between distribution functions. Figure 2.1 draws two examples of P, and marks the associated weights at ranks $u = .2$ and $u = .6$. The weight assigned to individuals at rank u equals the derivative of P at u. Note that the preference function must be concave and lie above the diagonal to ensure that W_P satisfies second-degree dominance.

Interpretation

A normative interpretation of the social welfare function defined by (2.1) can be made in terms of a theory for ranking distribution functions, as above, or as a value judgement of the trade-off between the mean and (in)equality in the distributions. By defining the ordering relation $\succeq$ on the set of Lorenz curves rather than on the set of distribution functions, Aaberge (2001) demonstrated that $\succeq$ can be represented by the following family of rank-dependent measures of inequality:

$$J_P(F) = 1 - \frac{1}{\mu_F} \int_0^1 P'(u) F^{-1}(u) du. \tag{2.2}$$

Following Ebert (1987), the social welfare function defined by (2.1) can then be expressed as

$$W_P(F) = \mu_F (1 - J_P(F)). \tag{2.3}$$

Equation (2.1) defines W_P as a weighted average of individual incomes where the weights decrease as a function of the individual's rank in the income distribution, while equation (2.3) shows directly how W_P reflects the trade-off between the mean and (in)equality in the distribu-

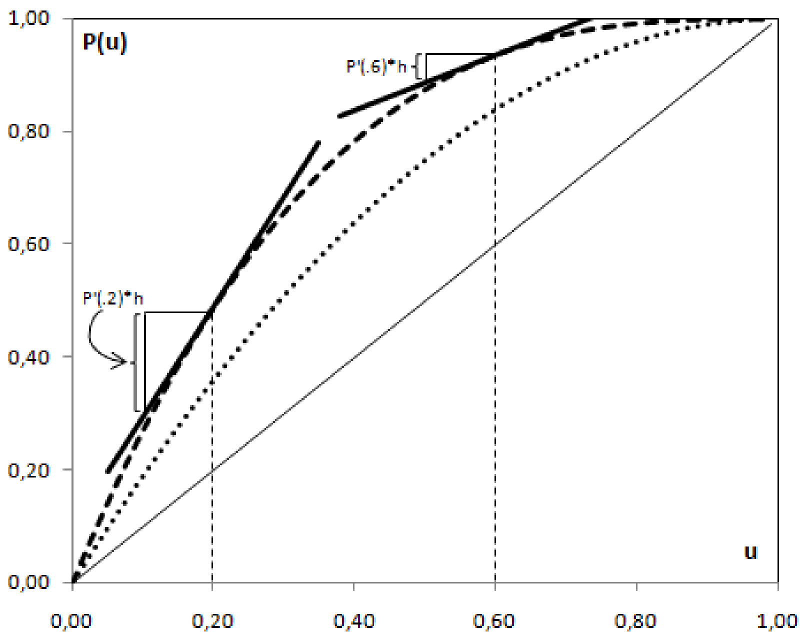


Figure 2.1: Examples of the preference function $P(\cdot)$ that preserves 3rd (dotted) and 4th degree (dashed) upward inverse stochastic dominance.
Note: The weight assigned to individuals at rank u equals the derivative of P at u.

tion of income. The product $\mu_F J_P(F)$ is a measure of the loss in social welfare due to inequality in the distribution of income. An inequality neutral planner would choose $P(t) = t$, which means that $W_P(F) = \mu_F$.

Parametric subfamilies

To quantify social welfare, it is necessary to work with parametric social welfare functions. The best known member of W_P is obtained by inserting for $P(t) = 2t - t^2$ in (2.2) and (2.3), in which case $J_P(F)$ is equal to the Gini coefficient and $W_P(F)$ is equal to the much used Gini social welfare function (see Sen, 1974). More generally, by choosing a parametric specification of P we can derive alternative parametric subfamilies of W_P.

If the preference function is defined by

$$P_{1k}(t) = 1 - (1-t)^{k-1}, k > 2 \tag{2.4}$$

then J_P becomes equal to the extended Gini family of inequality measures (Donaldson and Weymark, 1980) defined by

$$\begin{aligned} G_k(F) &= 1 - \frac{k-1}{\mu_F} \int_0^1 (1-t)^{k-2} F^{-1}(t) dt \\ &= \frac{1}{\mu_F} \int_0^\infty [1 - F(y)] \left[1 - (1 - F(y))^{k-2}\right] dx, \quad k > 2 \end{aligned} \tag{2.5}$$

where $G_3(F)$ is the Gini coefficient.[11] Inserting (2.5) in (2.3), W_P becomes equal to the extended Gini family of social welfare functions, defined by

$$W_{G_k}(F) = \int_0^\infty (1-F(y))^{k-1}\,dy = \mu_F\left[1-G_k(F)\right], \quad k>2 \tag{2.6}$$

If the preference function is instead defined by

$$P_{2k}(t) = \frac{(k-1)\,t - t^{k-1}}{k-2}, k>2 \tag{2.7}$$

then J_P becomes equal to the Lorenz family of inequality measures (Aaberge, 2000), defined by

$$\begin{aligned} D_k(F) &= 1-\frac{k-1}{(k-2)\mu_F}\int_0^1 (1-t^{k-2})F^{-1}(t)dt \\ &= \frac{1}{\mu_F\,(k-2)}\int_0^\infty F(x)\left(1-F^{k-2}(x)\right)dx, \quad k>2 \end{aligned} \tag{2.8}$$

where $D_3(F)$ is the Gini coefficient. Inserting (2.8) for $J_P(F)$ in (2.3), W_P becomes equal to the Lorenz family of social welfare functions

$$W_{D_k}(F) = \tfrac{k-1}{k-2}\mu_F - \tfrac{1}{k-2}\textstyle\int_0^\infty \left(1-F^{k-1}(x)\right)dx = \mu_F\left[1-D_k(F)\right], \quad k>2 \tag{2.9}$$

Since $\{\mu_F, W_{G_k}(F) : k=3,4,...\}$ and $\{\mu_F, W_{D_k}(F) : k=3,4,...\}$ uniquely determine the distribution function F (Aaberge, 2000), no information is lost by working directly with either of these parametric subfamilies and the mean.

2.2 Third-degree dominance and social welfare

When distribution functions intersect and second-degree dominance does not provide an unambiguous ranking of distribution functions, weaker criteria are required. This subsection considers third-degree inverse stochastic dominance and characterizes its relationship to W_P. We consider first the criterion of third-degree upward dominance, after which we introduce and analyze the criterion of third-degree downward dominance.

2.2.1 Upward dominance and social welfare

Let the function associated with second-degree inverse stochastic dominance be defined by

$$\Lambda_F^2(u) = \int_0^u F^{-1}(t)dt, \quad u\in[0,1] \tag{2.10}$$

[11] See Aaberge (2001) for an axiomatic justification for this family of inequality measures.

where the superscript 2 refers to inverse stochastic dominance of second-degree. To define third-degree upward inverse stochastic dominance, we use the notation

$$\Lambda_F^3(u) = \int_0^u \Lambda_F^2(t)dt = \int_0^u (u-t)F^{-1}(t)dt, \quad u \in [0,1] \tag{2.11}$$

where the second equality follows by inserting (2.10) in (2.11) and interchanging the order of integration.

Definition 2.2. A distribution F_1 is said to *third-degree upward inverse stochastic dominate* a distribution F_0 if and only if $\Lambda_{F_1}^3(u) \geq \Lambda_{F_0}^3(u)$ for all $u \in [0,1]$, and the inequality holds strictly for some $u \in (0,1)$.

From equation (2.11), it is clear that the criterion of third-degree upward dominance compares weighted sums of incomes, where the weights decrease linearly with the rank in the income distribution.

Interpretation

Equation (2.3) shows how W_P can be interpreted as reflecting the trade-off between the mean and (in)equality in the distribution of income. We now show that third-degree upward dominance has an analogous interpretation.

Let H be the conditional distribution function defined by $H(y) = Pr(Y \leq y \mid Y \leq F^{-1}(u)) = F(y)/u$, for any $y \leq F^{-1}(u)$. The quantile-specific lower tail mean is defined by

$$\mu_F(u) \quad = \mu_H = \int_0^{F^{-1}(u)} ydH(y) = \frac{\int_0^u F^{-1}(t)dt}{u} \tag{2.12}$$

and the quantile-specific lower tail Gini coefficient is defined by

$$G_3(u;F) \quad = \quad \frac{1}{\mu_H}\int_0^1 (2t-1)H^{-1}(t)dt = \frac{1}{u^2\mu_F(u)}\int_0^u (2t-u)F^{-1}(t)dt. \tag{2.13}$$

The quantile-specific lower tail Gini social welfare function is then given by $\mu_F(u)\,(1-G_3(u;F))$.

The following proposition shows that the criterion of third-degree upward dominance is equivalent to employing the Gini social welfare function to compare the welfare of individuals located in the lower tail of each quantile of the distributions.

Proposition 2.1. *Let F_1 and F_0 be members of $\mathscr{F}$. Then the following statements are equivalent:*
(i) F_1 third-degree upward inverse stochastic dominates F_0
(ii) $\mu_{F_1}(u)\,(1-G_3(u;F_1)) \geq \mu_{F_0}(u)\,(1-G_3(u;F_0))$ for all $u \in [0,1]$, *and the inequality holds strictly for some $u \in (0,1)$.*

Proof. This result follows by noting that

$$\Lambda_F^3(u) \quad = \quad \frac{u^2}{2}\mu_F(u)\left(1 - G_3(u;F)\right), \tag{2.14}$$

which is obtained by inserting (2.12) and (2.13) in (2.11). □

Transfer principle

To provide a normative justification for dominance criterion of third degree, more powerful principles than the Pigou-Dalton transfer principle are needed. To this end, Kolm (1976) introduced the principle of diminishing transfers, which for a fixed difference in income considers a transfer from a richer to a poorer person to be more equalizing the further down in the income distribution it takes place. As indicated by Shorrocks and Foster (1987) and Muliere and Scarsini (1989), the principle of diminishing transfers is, however, not consistent with third-degree upward inverse stochastic dominance. We will instead use an alternative version of the principle of diminishing transfers introduced by Mehran (1976) – and called the principle of positional transfer sensitivity by Zoli (1999) – to characterize third-degree upward inverse stochastic dominance.

In order to provide a formal definition of the principle of positional transfer sensitivity it will be useful to introduce the notation $\Delta_s W_P(\delta,h)$, which denotes the change in W_P of a fixed progressive transfer δ from an individual with rank $s+h$ to an individual with rank s. Further, let

$$\Delta_{st}^1 W_P(\delta,h) \equiv \Delta_s W_P(\delta,h) - \Delta_t W_P(\delta,h).$$

We can then define the principle of first-degree downside positional transfer sensitivity.

Definition 2.3. W_P satisfies the principle of first-degree downside positional transfer sensitivity (DPTS) if and only if $\Delta_{st}^1 W_P(\delta,h) > 0$, for all $s < t$.

To better understand first-degree DPTS and how it relates to the Pigou-Dalton transfer principle, consider Figure 2.2 where we draw the probability density of a right-skewed income distribution, denoted $f(x)$. We have also drawn two alternative transfers from richer to poorer, one from an individual at rank $t+h$ to an individual at rank t, and another from rank $s+h$ to rank s; the equal difference in rank h is reflected in the equal size of the shaded areas. Consider first the two transfers in isolation. According to the Pigou-Dalton transfer principle, both transfers should decrease inequality and hence increase welfare. According to first-degree DPTS, given that a fixed transfer takes place between two people with equal difference in ranks, the transfer at lower ranks has a stronger equalizing effect – and thus increases social welfare more – than the transfer at higher ranks. An inequality averse social planner who supports the principle of first-degree *DPTS* is said to exhibit downside positional inequality aversion of first-degree.

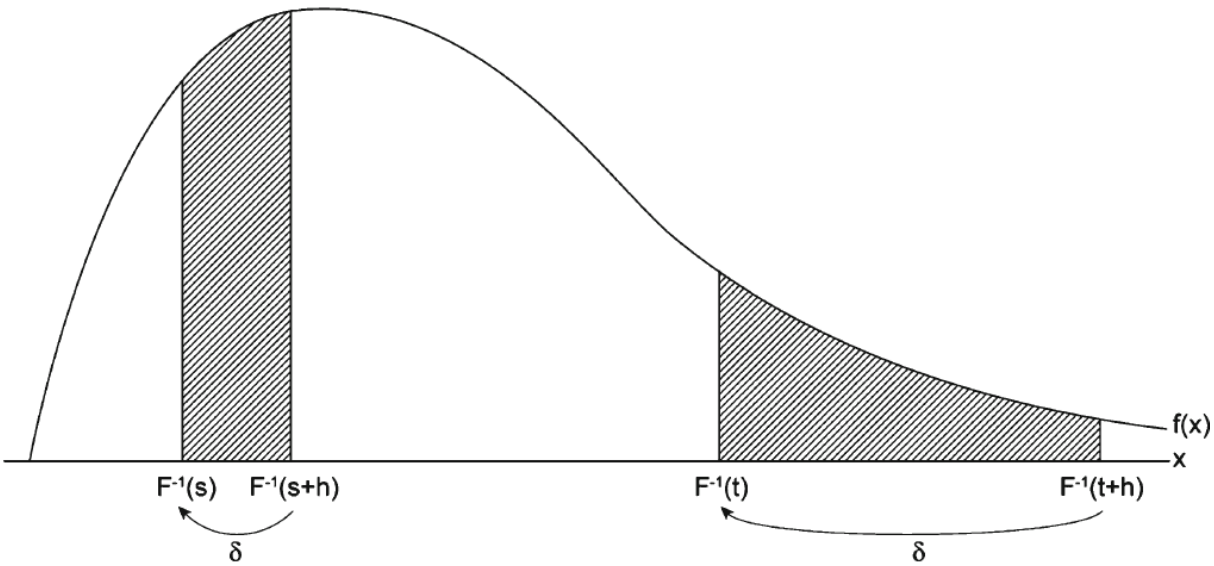


Figure 2.2: Income transfers and the principles of positional transfer sensitivity
Note: This figure draws the probability density of a right-skewed income distribution, denoted $f(x)$. We have also drawn two alternative transfers from richer to poorer, one from an individual at rank $t+h$ to an individual at rank t, and another from rank $s+h$ to rank s; the equal difference in rank h is reflected in the equal size of the shaded areas.

Equivalence result

Let $\mathscr{P}_3$ be the family of preference functions defined by

$$\mathscr{P}_3 = \left\{ P \in \mathscr{P} : P'''(t) > 0, \text{ for all } t \in (0,1) \, and \, P''(1) \leq 0 \right\} \tag{2.15}$$

The following result provides a characterization of the relationship between third-degree upward inverse stochastic dominance and the general family of welfare functions.

Theorem 2.1. *Let F_1 and F_0 be members of $\mathscr{F}$. Then the following statements are equivalent,*
(i) F_1 third-degree upward inverse stochastic dominates F_0
(ii) $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}_3$
(iii) $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}$ where W_P satisfies first-degree DPTS

Proof. In the appendix. □

The equivalence between (i) and (ii) in Theorem 2.1 reveals the least-restrictive set of social welfare functions that allows an unambiguous ranking of distribution functions in accordance with third-degree upward inverse stochastic dominance. This is ensured by imposing the requirement of a positive third-derivative on the preference function P. Further, the equivalence with (iii) provides a normative justification for ranking distribution functions according to third-degree upward dominance.[12]

[12]Mehran (1976) shows that J_P defined by (2.2) satisfies first-degree DPTS if and only if $P'''(t) > 0$, which is restated in the equivalence of (ii) and (iii) in Theorem 2.1. Aaberge (2000) demonstrates that J_P defined by (2.2)

2.2.2 Downward dominance and social welfare

Section 2.2.1 demonstrated that a social planner who supports the criterion of third-degree upward inverse stochastic dominance exhibits aversion to downside inequality. In some cases, however, the researcher may want ranking criteria that are more sensitive to income differences in the upper part of the distribution. One example is the growing literature on the long-run evolution of income distributions which devotes much attention to changes in top incomes (see e.g. Atkinson and Piketty, 2007; 2010).

To focus attention on differences in the upper part of the distribution, we introduce the criterion of third-degree *downward* inverse stochastic dominance. This criterion is obtained by aggregating the integrated inverse distribution function from above, rather than from below as in upward dominance. To define third-degree downward dominance, we use the notation

$$\tilde{\Lambda}_F^3(u) = \int_u^1 \Lambda_F^2(t)dt = (1-u)\mu - \int_u^1 (t-u)F^{-1}(t)dt, \quad u \in [0,1] \tag{2.16}$$

where the second equality follows from inserting (2.10) for Λ_F^2 and by interchanging the order of integration.

Definition 2.4. A distribution F_1 is said to *third-degree downward inverse stochastic dominate* a distribution F_0 if and only if $\tilde{\Lambda}_{F_1}^3(u) \geq \tilde{\Lambda}_{F_0}^3(u)$, for all $u \in [0,1]$, and the inequality holds strictly for some $u \in (0,1)$.

From equation (2.16), it is clear that the criterion of third-degree downward dominance compares the weighted sums of incomes, where the weights decrease linearly with the rank in the income distribution.

Interpretation

Equation (2.3) shows how W_P can be interpreted as reflecting the trade-off between the mean and (in)equality in the distribution of income. We now show that third-degree downward dominance has an analogous interpretation.

Let $\tilde{H}$ be the conditional distribution function defined by $\tilde{H}(y) = Pr(Y \leq y \mid Y \geq F^{-1}(u)) = (F(y)-u)/(1-u)$, for any $y \geq F^{-1}(u)$. The quantile-specific upper tail mean is defined by

$$\tilde{\mu}_F(u) \quad = \mu_{\tilde{H}} = \quad \int_{F^{-1}(u)}^1 y d\tilde{H}(y) = \frac{\int_u^1 F^{-1}(t)dt}{1-u} \tag{2.17}$$

satisfies the principle of diminishing transfers under conditions that depend on both the functional form of the preference function P and the shape of the income distribution F.

and the quantile-specific upper tail Gini coefficient is defined by

$$D_3(u;F) \quad = \quad \frac{1}{\mu_{\tilde{H}}}\int_0^1 (2t-1)\tilde{H}^{-1}(t)dt = \frac{\int_u^1 (2t-u-1)F^{-1}(t)dt}{(1-u)^2\tilde{\mu}_F(u)}. \qquad (2.18)$$

The quantile-specific upper tail Gini social welfare function is then given by $\tilde{\mu}_F(u)\,(1-D_3(u;F))$.

The following proposition shows that the criterion of third-degree downward dominance is a sequential comparison of a weighted sum of the mean income of the poorest u percent, and the Gini social welfare of the richest $(1-u)$ percent of the population.

Proposition 2.2. *Let F_1 and F_0 be members of $\mathscr{F}$. Then the following statements are equivalent:*
(i) F_1 third-degree downward inverse stochastic dominates F_0
(ii) $u\mu_{F_1}(u)+\frac{(1-u)}{2}\tilde{\mu}_{F_1}(u)\,(1-D_3(u;F_1)) \geq u\mu_{F_0}(u)+\frac{(1-u)}{2}\tilde{\mu}_{F_0}(u)\,(1-D_3(u;F_0))$ for all $u\in[0,1]$ *and the inequality holds strictly for some $u\in(0,1)$.*

Proof. This result is obtained by noting that

$$\tilde{\Lambda}_F^3(u) \quad = \quad u(1-u)\mu_F(u)+\frac{(1-u)^2}{2}\tilde{\mu}_F(u)\,(1-D_3(u;F))\,, \qquad (2.19)$$

which follows by inserting (2.17) and (2.18) in (2.16). □

Transfer principle

To provide a normative justification for downward dominance of third degree, more powerful principles than the Pigou-Dalton transfer principle are needed. We will employ the principle of *upside* positional transfer sensitivity – introduced by Aaberge (2009) for analyzing Lorenz dominance – to characterize third-degree downward inverse stochastic dominance.

As above, let $\Delta_s W_P(\delta,h)$ denote the change in W_P of a fixed progressive transfer δ from an individual with rank $s+h$ to an individual with rank s, and let

$$\Delta_{st}^1 W_P(\delta,h) \equiv \Delta_s W_P(\delta,h) - \Delta_t W_P(\delta,h).$$

We can then define the principle of first-degree upside positional transfer sensitivity.

Definition 2.5. W_P satisfies the principle of first-degree upside positional transfer sensitivity (UPTS) if and only if $\Delta_{st}^1 W_P(\delta,h) < 0$, for all $s<t$.

To better understand first-degree UPTS and how it relates to the Pigou-Dalton transfer principle and first-degree DPTS, revisit Figure 2.2. We have drawn two alternative transfers from richer to poorer: One from an individual at rank $t+h$ to an individual at rank t, and another from rank $s+h$ to rank s; the equal difference in rank h is reflected in the equal size of the shaded areas. This implies that the number of people between the donor and the receiver is the same.

Consider first the two transfers in isolation. According to the Pigou-Dalton transfer principle, both transfers should decrease inequality and hence increase welfare. According to first-degree UPTS, given that a fixed transfer takes place between two persons with equal difference in ranks, the transfer at lower ranks has a weaker equalizing effect – and thus increases social welfare less – than the transfer at higher ranks. An inequality averse social planner that supports the principle of first-degree *UPTS* is therefore said to exhibit upside positional inequality aversion of first-degree. The choice between DPTS and UPTS clarifies, therefore, whether equalizing transfers between poorer individuals should be considered more or less important for social welfare as compared to equalizing transfers between richer individuals.

Equivalence result

Let $\tilde{\mathscr{P}}_3$ be the family of preference functions defined by

$$\tilde{\mathscr{P}}_3 = \left\{ P \in \mathscr{P} : P'''(t) < 0 \, for \, all \, t \in (0,1) \, and \, P''(0) \leq 0 \right\}. \tag{2.20}$$

The following result provides a characterization of the relationship between third-degree downward inverse stochastic dominance and the general family of welfare functions.

Theorem 2.2. *Let F_1 and F_0 be members of $\mathscr{F}$. Then the following statements are equivalent,*
(i) F_1 third-degree downward inverse stochastic dominates F_0
(ii) $W_P(F_1) > W_P(F_0)$ for all $P \in \tilde{\mathscr{P}}_3$
(iii) $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}$ where W_P satisfies first-degree UPTS

Proof. In the appendix. □

The equivalence between (i) and (ii) in Theorem 2.2 reveals the least-restrictive set of social welfare functions that allows an unambiguous ranking of distribution functions in accordance with third-degree downward inverse stochastic dominance. This is ensured by imposing the requirement of a negative third-derivative on the preference function P. Further, the equivalence with (iii) provides a normative justification for ranking distribution functions according to third-degree downward dominance. By comparing (iii) in Theorems 2.1 and 2.2, it is clear that the choice between third-degree upward dominance and third-degree downward dominance depends on whether income differences between poorer individuals are viewed as more or less important for social welfare as compared to income differences between richer individuals.

2.3 Dominance of *i*th-degree and social welfare

In some cases, neither upward nor downward dominance of third-degree allows an unambiguous ranking of the distribution functions under comparison. This subsection therefore introduces the full hierarchical sequences of nested inverse stochastic dominance criteria, allowing ranking of any set of distribution functions. We further characterize the relationship between W_p and upward or downward dominance of any degree.

To define upward inverse stochastic dominance of degree i, we use the notation

$$\begin{aligned}\Lambda_F^i(u) &= \int_0^u \Lambda_F^{i-1}(t)dt = \frac{1}{(i-3)!}\int_0^u (u-t)^{i-3}\Lambda_F^2(t)dt \\ &= \frac{1}{(i-2)!}\int_0^u (u-t)^{i-2}F^{-1}(t)dt,\ i=3,4,... \end{aligned} \tag{2.21}$$

To define downward inverse stochastic dominance of degree i, we use the notation

$$\begin{aligned}\tilde{\Lambda}_F^i(u) &= \int_u^1 \tilde{\Lambda}_F^{i-1}(t)dt = \frac{1}{(i-3)!}\int_u^1 (t-u)^{i-3}\Lambda_F^2(t)dt \\ &= \frac{1}{(i-2)!}\left[(1-u)^{i-2}\mu_F - \int_u^1 (t-u)^{i-2}F^{-1}(t)dt\right]\ i=3,4,... \end{aligned} \tag{2.22}$$

Definition 2.6. A distribution F_1 is said to ith-degree upward inverse stochastic dominate F_0 if and only if $\Lambda_{F_1}^i(u) \geq \Lambda_{F_0}^i(u)$, for all $u \in [0,1]$, and the inequality holds strictly for some $u \in (0,1)$.

Definition 2.7. A distribution F_1 is said to ith-degree downward inverse stochastic dominate F_0 if and only if $\tilde{\Lambda}_{F_1}^i(u) \geq \tilde{\Lambda}_{F_0}^i(u)$, for all $u \in [0,1]$, and the inequality holds strictly for some $u \in (0,1)$.

From equation (2.21) and (2.22), it is clear that the criteria of both ith degree upward and downward dominance compare the weighted sums of incomes, where the weights decrease with the rank in the income distribution.[13] As will be demonstrated below, however, the choice between higher degree of upward and downward dominance clarifies whether preferences of the social planner gives priority to reduction of inequality in the lower or the upper part of the income distribution.

Interpretation

We now show that upward and downward dominance of degree i can be interpreted as reflecting trade-offs between the mean and (in)equality in the distribution of income. To this end, we employ the two parametric subfamilies of W_p presented above: The first is the extended Gini family of social welfare functions $W_{G_k}(F)$, defined by equation (2.6); the second is the Lorenz family of social welfare functions $W_{D_k}(F)$, defined by equation (2.9).

[13] Note that Definitions 2.6 and 2.7 do not require any restrictions on the distribution functions and thus are less restrictive than the definitions of stochastic dominance proposed by Whitmore (1970) and Chew (1983).

The quantile-specific lower tail extended Gini family of inequality measures is defined by

$$G_i(u;F) = 1 - \frac{i-1}{\mu_H}\int_0^1 (1-t)^{i-2}H^{-1}(t)dt = 1 - \frac{i-1}{u^{i-1}\mu_F(u)}\int_0^u (u-t)^{i-2}F^{-1}(t)dt, \quad (2.23)$$

and the associated quantile-specific lower tail extended Gini family of social welfare functions can then be expressed as $\mu_F(u)(1-G_i(u;F))$.

Similarly, the quantile-specific upper tail Lorenz family of inequality measures is defined by

$$\begin{aligned} D_i(u;F) &= 1 - \frac{i-1}{(i-2)\mu_{\tilde{H}}}\int_0^1 (1-t^{i-2})\tilde{H}^{-1}(t)dt \quad (2.24) \\ &= 1 - \frac{i-1}{(i-2)(1-u)^{i-1}\tilde{\mu}_F(u)}\int_u^1 \left[(1-u)^{i-2} - (t-u)^{i-2}\right]F^{-1}(t)dt, \end{aligned}$$

and the associated quantile-specific upper tail Lorenz family of social welfare functions can then be expressed as $\tilde{\mu}_F(u)(1-D_i(u;F))$.

Proposition 2.3 shows that the criterion of i^{th}-degree upward dominance is equivalent to employing the Gini social welfare function of order i to compare welfare among individuals located at the lower tail of each quantile of the distributions. Proposition 2.4 shows that the criterion of third degree downward dominance corresponds to a sequential comparison of a weighted sum of the mean income of the poorest u percent, and the social welfare of the richest $(1-u)$ percent of the population according to the Lorenz social welfare function of order i.

Proposition 2.3. *Let F_0 and F_1 be members of $\mathscr{F}$. Then the following statements are equivalent:*
(i) F_1 ith-degree upward inverse stochastic dominates F_0
(ii) $\mu_{F_1}(u)(1-G_i(u;F_1)) \geq \mu_{F_0}(u)(1-G_i(u;F_0))$ for all $u \in [0,1]$, *and the inequality holds strictly for some $u \in (0,1)$.*

Proof. This result is obtained by noting that

$$\Lambda_F^i(u) = \frac{u^{i-1}}{(i-1)!}\mu_F(u)(1-G_i(u;F)), \quad (2.25)$$

which follows by inserting (2.12) and (2.23) in (2.21). □

Proposition 2.4. *Let F_0 and F_1 be members of $\mathscr{F}$. Then the following statements are equivalent:*
(i) F_1 ith-degree downward inverse stochastic dominates F_0
(ii) $u\mu_{F_1}(u) + \frac{(i-2)}{(i-1)}(1-u)\tilde{\mu}_{F_1}(u)(1-D_i(u;F_1)) \geq u\mu_{F_0}(u) - \frac{(i-2)}{(i-1)}(1-u)\tilde{\mu}_{F_0}(u)(1-D_i(u;F_0))$
for all $u \in [0,1]$, *and the inequality holds strictly for some $u \in (0,1)$.*

Proof. This result is obtained by noting that

$$\tilde{\Lambda}_F^i(u) = \frac{u(1-u)^{i-2}}{(i-2)!}\mu_F(u) + \frac{(i-2)(1-u)^{i-1}}{(i-1)!}\tilde{\mu}_F(u)\left(1 - D_i(u;F)\right), \tag{2.26}$$

which follows by inserting (2.17) and (2.24) in (2.22). □

Transfer principles

To provide a normative justification for upward (downward) dominance of degree i, we employ generalizations of the principle of downside (upside) positional transfer sensitivity. As above, let $\Delta_s W_P(\delta,h)$ denote the change in W_P of a fixed progressive transfer δ from an individual with rank $s+h$ to an individual with rank s, and let $\Delta_{st}^1 W_P(\delta,h) = \Delta_s W_P(\delta,h) - \Delta_t W_P(\delta,h)$. Further, let

$$\Delta_{st}^i W_P(\delta,h_1,h_2,\ldots,h_i) \equiv \Delta_{st}^{i-1} W_P(\delta,h_1,h_2,\ldots,h_{i-1}) - \Delta_{s+h_i,t+h_i}^{i-1} W_P(\delta,h_1,h_2,\ldots,h_{i-1}), \tag{2.27}$$

for $i = 2,3,\ldots$, denote the difference in the change in social welfare from a series of progressive transfers at lower ranks (s) compared to higher ranks (t) in the income distribution. We can then define the principles of downside and upside positional transfer sensitivity of degree i.

Definition 2.8. W_P satisfies the principle of downside positional transfer sensitivity (DPTS) of degree i if and only if, for all $k = 1,2,\ldots,i$

$$(-1)^k \Delta_{st}^k W_P(\delta,h) > 0, \quad \text{when } s < t.$$

Definition 2.9. W_P satisfies the principle of upside positional transfer sensitivity (UPTS) of degree i if and only if, for all $k = 1,2,\ldots,i$

$$\Delta_{st}^k W_P(\delta,h) > 0, \quad \text{when } s < t.$$

Given two alternative sequences of fixed transfers which take place between people with equal difference in ranks, ith degree UPTS (DPTS) states that the sequence of transfers at lower ranks have a stronger (weaker) equalizing effect – and thus increase social welfare more (less) – than the sequence of transfers at higher ranks. Further, a social planner that supports the principle of ith degree UPTS (DPTS) exhibits relatively higher inequality aversion in the lower (upper) parts of the distribution, as compared to a social planner that supports the principle of $(i-1)$th-degree UPTS (DPTS). An inequality averse social planner that supports the principle of ith-degree UPTS (DPTS) is therefore said to exhibit downside (upside) positional inequality aversion of degree i.[14] Since UPTS (DPTS) of degree i are stronger criteria than UPTS (DPTS)

[14]Note that ith-degree UPTS can be considered as an alternative to the ith-degree transfer principle introduced by Fishburn and Willig (1984) as an extension of Kolm's principle of diminishing transfers.

of degree $i-1$, it seems natural that a social planner who supports the latter will also support the former.

Equivalence result

Let $P^{(j)}$ denote the jth-degree derivative of P. The family of preference functions $\mathscr{P}_i$ is defined by

$$\begin{aligned}\mathscr{P}_i \;=\; \Big\{P \in \mathscr{P} : (-1)^{i-1}P^{(i)}(t) > 0 \\ \text{and } (-1)^{j-1}P^{(j)}(1) \geq 0 \text{ for all } j = 2,3,\ldots,i-1\Big\}\end{aligned} \tag{2.28}$$

while the family of preference functions $\tilde{\mathscr{P}}_i$ is defined by

$$\begin{aligned}\tilde{\mathscr{P}}_i \;=\; \Big\{P \in \mathscr{P} : P^{(i)}(t) < 0 \\ \text{and } P^{(j)}(0) \leq 0 \text{ for all } j = 2,3,\ldots,i-1\Big\}\end{aligned} \tag{2.29}$$

The following theorems provide a characterization of the relationship between ith-degree upward and downward inverse stochastic dominance and the general family of welfare functions.

Theorem 2.3. *Let F_1 and F_0 be members of $\mathscr{F}$. Then for $i = 3,4,...$ the following statements are equivalent,*
(i) F_1 ith-degree upward inverse stochastic dominates F_0
(ii) $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}_i$
(iii) $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}$ where W_P satisfies DPTS of degree $i-2$

Proof. In the appendix. □

Theorem 2.4. *Let F_1 and F_0 be members of $\mathscr{F}$. Then for $i = 3,4,\ldots$ the following statements are equivalent*
(i) F_1 ith-degree downward inverse stochastic dominates F_0
(ii) $W_P(F_1) > W_P(F_0)$ for all $P \in \tilde{\mathscr{P}}_i$
(iii) $W_P(F_1) > W_P(F_0)$ for all $P \in \mathscr{P}$ where W_P satisfies UPTS of degree $i-2$

Proof. In the appendix. □

The equivalence between (i) and (ii) in Theorems 2.3 and 2.4 reveals the least-restrictive set of social welfare functions that allows an unambiguous ranking of distribution functions in accordance with ith degree upward or downward inverse stochastic dominance.

Upward dominance of degree i is ensured by imposing positive (negative) ith-degree derivative if i is odd (even) on the preference function P. Together with the boundary condition, this makes sure that the implied set of weights becomes more progressive as i increases. This means that a social planner who employs the criterion of ith-degree upward dominance pays more attention to inequality in the lower than in the upper part of the income distribution as compared to a social planner who employs the criterion of $(i-1)$th-degree upward dominance.

Downward dominance of degree i is ensured by imposing negative ith-degree derivative on the preference function P. Together with the boundary condition, this makes sure that the implied set of weights becomes more progressive as i increases. This means that a social planner who employs the criterion of ith-degree downward dominance pays more attention to inequality in the upper than in the lower part of the income distribution as compared to a social planner who employs the criterion of $(i-1)$th-degree downward dominance.

The equivalence between (i) and (iii) in Theorems 2.3 and 2.4 provides normative justification for ranking distribution functions according to ith-degree upward and downward dominance. By comparing (iii) in these two theorems, it is clear that the choice between ith-degree upward dominance and ith-degree downward dominance depends on whether income differences between poorer individuals are viewed as more or less important for social welfare as compared to income differences between richer individuals.

Remark. The dominance relations are transitive. To see this, assume
(i) F_1 ith-degree upward (downward) dominates F_2
(ii) F_2 $(i-k)$th degree upward (downward) dominates F_3.
For $k=0$, it follows from Definitions 2.6 and 2.7 that (i) and (ii) imply that F_1 ith-degree upward (downward) inverse stochastic dominates F_3.
From Equations (2.21) and (2.22), it follows that $\Lambda_{F_1}^{i-1}(u) \geq \Lambda_{F_2}^{i-1}(u)$ for all u implies $\Lambda_{F_1}^{i}(u) \geq \Lambda_{F_2}^{i}(u)$ for all u. For $k=0,1,\ldots$, (i) and (ii) therefore imply that F_1 $(i-k)$th-degree upward (downward) inverse stochastic dominates F_3.

2.4 The limits of the dominance criteria

The proposed sequences of dominance criteria along with Theorems 2.3 and 2.4 suggest two complementary strategies for successively narrowing the general family of social welfare functions in order to unambiguously rank any set of distribution functions. Though the theorems are only valid for finite i, to understand their normative implications it is helpful to consider the limits of the sequences of dominance criteria.

As $i \to \infty$ we get from equations (2.21) and (2.22)

$$(i-1)!\Lambda^i(u) \quad \to \quad \begin{cases} 0, & 0 \leq u < 1 \\ F^{-1}(0+), & u = 1 \end{cases} \tag{2.30}$$

$$(i-2)!\tilde{\Lambda}^i(u) \quad \to \quad \begin{cases} \mu_F, & u = 0 \\ 0, & 0 < u \leq 1 \end{cases} \tag{2.31}$$

where $F^{-1}(0+)$ denotes the lowest income in F. In the limit, upward and downward inverse stochastic dominance therefore depend only on the income of the worst-off income recipient and the average income, respectively.

The highest degree of downside inequality aversion is achieved when focus is exclusively turned to the situation of the poorest in the population. In this case the social welfare function corresponds to *the Rawlsian maximin criterion*. By contrast, the highest degree of upside inequality aversion is achieved when focus is exclusively turned to the mean income. In this case, the social welfare function corresponds to the *utilitarian criterion*. The utilitarian criterion is "dual" to the Rawlsian maximin criterion in the sense that it is compatible with the limiting case of downward inverse stochastic dominance. When the comparison of distribution functions is based on the utilitarian criterion, the distribution function for which the mean income is largest is preferred, regardless of all other differences.

3 Inverse stochastic dominance and parametric families of social welfare functions

Until now, the results and discussion have centered on characterizing the relationship between inverse stochastic dominance criteria and W_P in the ranking of intersecting distribution functions. This section extends our framework to not only rank distributions, but also quantify the social welfare level of a dominating distribution as compared to a dominated distribution. To this end, we employ the two parametric subfamilies of W_p presented above: The first is the extended Gini family of social welfare functions $W_{G_k}(F)$, defined by equation (2.6); the second is the extended Lorenz family of social welfare functions $W_{D_k}(F)$, defined by equation (2.9). Since $\{\mu_F, W_{G_i}(F) : i = 3,4,...\}$ and $\{\mu_F, W_{D_i}(F) : i = 3,4,...\}$ uniquely determine the distribution function F (Aaberge, 2000), no information is lost by working directly with either of these parametric subfamilies and the mean.

Upward dominance and the extended Gini family

Corollary 3.1 sorts the members of the Gini family of social welfare functions into subfamilies according to their relationship to upward inverse stochastic dominance. This allows us to identify the largest subfamily of $W_{G_i}(F)$ that ranks consistently with upward dominance of a given degree, and quantify the social welfare level of the dominating distribution as compared to the dominated distribution. From Theorem 2.3, we get the following result.

Corollary 3.1. *Let F_1 and F_0 be members of $\mathscr{F}$. Then for $i = 2,3\ldots$*
(i) F_1 i^{th} degree upward inverse stochastic dominates F_0
implies
(ii) $W_{G_k}(F_1) > W_{G_k}(F_0)$ for $k > i$

Remark. The extended Gini family of social welfare functions has the following properties,
(i) W_{G_i} obeys the Pigou-Dalton principle of transfers for $i > 2$.

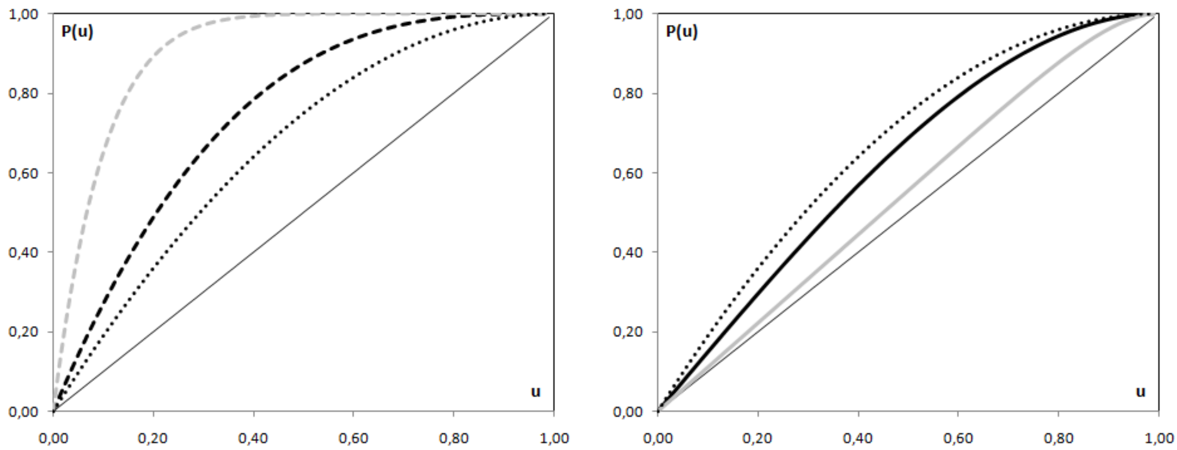


Figure 3.1: Examples of the preference function P that preserves 2nd, 3rd and 10th degree inverse stochastic dominance, upwards (left panel) and downwards (right panel).
Note: The weight assigned to individuals at rank u equal the derivative of P at u. The parametric forms of P are defined in Section 2.1.

(ii) W_{G_i} obeys the principles of DPTS up to and including $(i-2)$th-degree for $i = 3, 4, \ldots$.
(iii) The sequence $\{W_{G_i}\}$approaches μ_F when $i \to 2$
(iv) The sequence $\{W_{G_i}\}$ approaches the Rawlsian maxi-min criterion when $i \to \infty$.

The left panel of Figure 3.1 displays the preference function $P_{1k}(t)$ defined by (2.4) when $k = 3$, $k = 4$ and $k = 10$. As we increase the degree of upward dominance preserved by W_{G_k}, we see how the preference function becomes more sensitive to income differences in the lower part of the distribution. This is also illustrated in Panel (a) of Table 3.1. This table shows how $P_{1k}(t)$ assigns weights to incomes at selected quantiles relative to the weight assigned to the median income, both when $k = 3, 4, 5, 6$ and in the limits as $k \to 2$ and $k \to \infty$. The highest degree of downside inequality aversion occurs as $k \to \infty$, which corresponds to the Rawlsian maximin criterion. At the other extreme, $k \to 2$ and W_{G_k} equals the mean income.

Downward dominance and the extended Lorenz family

Corollary 3.2 sorts the members of the Lorenz family of social welfare functions into subfamilies according to their relationship to downward inverse stochastic dominance. This allows us to identify the largest subfamily of $W_{D_i}(F)$ that ranks consistently with downwards dominance of a given degree, and quantify the social welfare level of the dominating distribution as compared to the dominated distribution. From Theorem 2.4, we get the following result.

Corollary 3.2. *Let F_1 and F_0 be members of $\mathscr{F}$. Then for $i = 2, 3, \ldots$*
(i) F_1 ith degree downward inverse stochastic dominates F_0
implies
(ii) $W_{D_k}(F_1) > W_{D_k}(F_0)$ for $k > i$

Table 3.1: Weights in W_{G_k} and W_{D_k} at selected quantiles relative to the weight at the median

	Quantile					
	.01	*.05*	*.30*	*.70*	*.95*	*.99*
Panel (a): Gini social welfare function (upward)						
$k=2$	1.00	1.00	1.00	1.00	1.00	1.00
$k=3$	2.00	1.90	1.40	0.60	0.10	0+
$k=4$	4.00	3.61	1.96	0.36	0.01	0+
$k=5$	8.00	6.86	2.74	0.22	0.00	0+
$k=6$	16.00	13.03	3.84	0.13	0.00	0+
$k\to\infty$	∞	0	0	0	0	0
Panel (b): Lorenz social welfare function (downward)						
$k=3$	2.00	1.90	1.40	0.60	0.10	0+
$k=4$	1.33	1.33	1.21	0.68	0.13	0+
$k=5$	1.14	1.14	1.11	0.75	0.16	0+
$k=6$	1.07	1.07	1.06	0.81	0.20	0+
$k\to\infty$	1	1	1	1	1	1-

Note: The parametric forms of the weighting function P are defined in Section 2.1.

Remark. The extended Lorenz family of social welfare functions has the following properties,
(i) W_{D_i} obeys the Pigou-Dalton principle of transfers for $i>2$.
(ii) W_{D_i} obeys the principles of UPTS up to and including $(i-2)$th-degree for $i=3,4,\ldots$.
(iii) The sequence $\{W_{D_i}\}$ approaches the Bonferroni welfare function $\int[1-F(x)(1-\log F(x))]\,dx$ when $i\to 2$
(iv) The sequence $\{W_{D_i}\}$ approaches μ_F as $i\to\infty$
(v) The sequence $\{i(W_{D_i}-\mu_F)\}$ approaches $-F^{-1}(1-)$ as $i\to\infty$, which means that the distribution with the lowest maximum income is considered preferable provided that the distributions in question have equal mean income.

The right panel of Figure 3.1 displays the preference function $P_{2k}(t)$ when $k=3$, $k=4$ and $k=10$. As we increase the degree of downward dominance preserved by W_{D_i}, we see how the preference function becomes more sensitive to income differences in the upper part of the distribution. This is also illustrated in Panel (b) of Table 3.1. This table shows how $P_{2k}(t)$ assigns weights to incomes at selected quantiles relative to the weight assigned to the median income, both when $k=3,4,5,6$ and at the limit when $k\to\infty$. The highest degree of upside inequality aversion occurs as $k\to\infty$, which corresponds to the utilitarian criterion.

4 Asymptotic theory

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

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

Asymptotic distribution theory

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

$$Q_n(u) = \sqrt{n}\left(F_n^{-1}(u) - F^{-1}(u)\right) \tag{4.1}$$

Approximations to the variances of $\Lambda^i_{F_n}(u)$ and $\tilde{\Lambda}^i_{F_n}(u)$ and the asymptotic properties of $\Lambda^i_{F_n}(u)$ and $\tilde{\Lambda}^i_{F_n}(u)$ can be obtained by considering the limiting distribution of the empirical processes

$Y_n^i(u)$ and $\tilde{Y}_n^i(u)$ defined by

$$Y_n^i(u) = \sqrt{n}\left[\Lambda_{F_n}^i(u) - \Lambda_F^i(u)\right] = \frac{1}{(i-2)!}\int_0^u (u-t)^{i-2} Q_n(t)\,dt \tag{4.2}$$

and

$$\tilde{Y}_n^i(u) = \sqrt{n}\left[\tilde{\Lambda}_{F_n}^i(u) - \tilde{\Lambda}_F^i(u)\right] = \frac{1}{(i-2)!}\left[(1-u)^{i-2}\int_0^1 Q_n(t)\,dt - \int_u^1 (t-u)^{i-2} Q_n(t)\,dt\right] \tag{4.3}$$

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 (4.2) and (4.3) it is convenient to consider the empirical process

$$V_n(u) = \int_{a(u)}^{b(u)} w(u,t)\, Q_n(t)\,dt \tag{4.4}$$

which suggests that it will be useful to start with the process $Q_n(u)$ defined in (4.1).

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

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

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 4.2. *Suppose the conditions of Theorem 4.1 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. In the appendix. □

The following result is obtained from Theorems 4.1 and 4.2 by inserting $a(u) = 0$, $b(u) = u$ and $w(u,t) = (u-t)^{i-2} / (i-2)!$ in expression (4.4).

Corollary 4.1. *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 4.2. *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. In the appendix. □

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

$$2\sum_{j=1}^{\infty}\frac{sin\,(j\pi s)\,sin\,(j\pi t)}{(j\pi)^2} = s\,(1-t)\,,\, 0\le s\le t\le 1, \tag{4.5}$$

we get as an immediate consequence of Corollary 4.1 the following result.

Corollary 4.3. *Under the conditions of Theorem 4.1, $Y_n^i(u)$ has asymptotic covariance function given by*

$$\begin{aligned} v_i^2(u,v) &= \sum_{j=1}^{\infty} h_j^i(u)\,h_j^i(v) \qquad (4.6)\\ &= \frac{1}{[(i-2)!]^2}\left\{2\int_{F^{-1}(0)}^{F^{-1}(u)}\int_{F^{-1}(0)}^{y}[(u-F(x))\,(v-F(y))]^{i-2}F(x)\,(1-F(y))\,dxdy\right.\\ &\quad\left.+\int_{F^{-1}(u)}^{F^{-1}(v)}\int_{F^{-1}(0)}^{F^{-1}(u)}[(u-F(x))\,(v-F(y))]^{i-2}F(x)\,(1-F(y))\,dxdy\right\} \end{aligned}$$

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 4.3, we get the following result from Corollary 4.2 by applying Fubini's theorem (e.g. Royden, 1963) and the identity (4.5).

Corollary 4.4. *Under the conditions of Theorem 4.1, $\tilde{Y}_n^i(u)$ has asymptotic covariance function given by*

$$\begin{aligned} \eta_i^2(u,v) &= \sum_{j=1}^{\infty}\tilde{h}_j^i(u)\,\tilde{h}_j^i(v) \qquad (4.7)\\ &= 2\,[(1-u)\,(1-v)]^{i-2}\,\lambda_{i22}(0,0)-(1-u)^{i-2}\,[\lambda_{i2i}(u,v)+\lambda_{ii2}(u,v)+\gamma_{i2i}(0,v)]\\ &\quad-(1-v)^{i-2}\,[\lambda_{i2i}(u,u)+\lambda_{ii2}(u,u)+\gamma_{i2i}(0,u)]+\left[\lambda_{iii}(u,v)+\tilde{\lambda}_{iii}(u,v)+\gamma_{iii}(u,v)\right] \end{aligned}$$

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 4.1 and Corollary 4.2, 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 (4.6), and the distribution of

$$\sqrt{n}\frac{\tilde{\Lambda}_{F_n}^i(u) - \tilde{\Lambda}_F^i(u)}{\eta_i(u,u)}$$

tends to the $N(0,1)$-distribution for fixed u, where $\eta_i^2(u,u)$ is given by (4.7).

Confidence intervals and bands

To get an idea of how reliable $\Lambda_{F_n}^i(u)$ and $\tilde{\Lambda}_{F_n}^i(u)$ are as estimates of $\Lambda_F^i(u)$ and $\tilde{\Lambda}_F^i(u)$, we have to construct confidence bands based on $\Lambda_{F_n}^i(u)$ and $\tilde{\Lambda}_{F_n}^i(u)$, 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 4.1 and 4.2 and Billingsley (1968, Th. 5.1), we find that K_n converges in distribution to

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

We use the following notation.

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

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

In the empirical analysis, we apply the distribution theory to test for upward and downward inverse stochastic dominance. We start by estimating 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. Under the null hypothesis of $F_1^{-1}(u) = F_0^{-1}(u)$ for all $u \in [0,1]$, we can estimate $h_j(u)$ and $\tilde{h}_j(u)$ by mixing the distributions and use a kernel estimate of the density function with an epanechnikov kernel. To calculate $T_m(u)$ we take 1000 random normal draws, set $\psi = 1$, and then calculate K_m^* taking the maximum of $|T_m(u)|$. Finally, we repeat this procedure 1000 times, and calculate the critical value at confidence level p as the p percentile of the simulated distribution of K_m^*.

5 Empirical applications

5.1 Distribution of household income in booms and busts

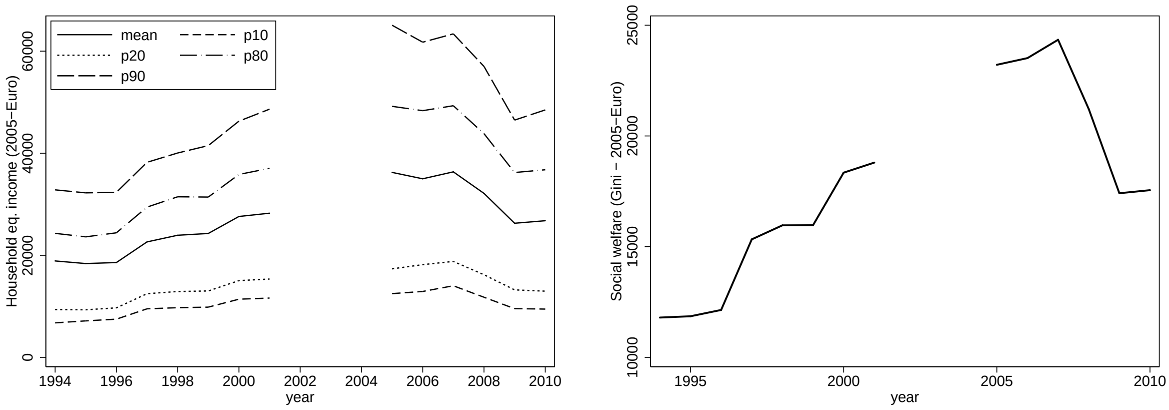
A large body of evidence suggests that inequality growth in the UK over the past few decades has been episodic and strongly related to the business cycle (see e.g. Blundell and Etheridge, 2010). From 1993 onwards, the economy moved out of a recession and into a period of stable and moderate income growth across most of the income distribution. Then, from the late 1990s, a further rise in income occurred, largely concentrated at the upper part of the income distribution. The recession that followed the financial crisis in 2007/2008 led to sharp falls in incomes, especially at the upper part of the income distribution.

Our framework can be used to make unambiguous statements about the social welfare implications of these changes in the household income distribution. Our data come from the European Community Household Panel (ECHP) for 1995–2001, and from the European Union Statistics on Income and Living Conditions (EU-SILC) for 2005–2010.[15] In each year, we restrict the sample to households with a male aged 25-64. We focus on the distribution of individual equivalent income, after adjusting for inflation and differences in household size and composition.[16] Using our data, Panel (a) of Figure 5.1 displays the evolution at different parts of the equivalent income distribution.

To assess the changes in the distribution of individual equivalent income, we make pairwise yearly comparisons of all the distributions. Table 5.1 shows the ranking on the basis of second-degree dominance, denoting by ">" if the earlier year dominates, and by "<" if the later year dominates. We can see that 45 of a possible 91 pairwise yearly comparisons can be ranked on the basis of second-degree dominance. Furthermore, all but eight of these rankings are statistically significant at conventional levels. This still leaves us a long way short of a complete

[15] Unfortunately, these datasets do not provide information on income for the years 2002–2004.

[16] To adjust for differences in household size and composition, we use the OECD equivalence scale.



(a) Means and percentiles of individual equiv. income (b) Social welfare, per capita

Figure 5.1: Time trends in the UK distribution of income and social welfare

Note: The data come from ECHP (1994–2001) and EU-SILC (2005–2010). We use the population weights supplied by EU-Stat. In Panel (a), we show the evolution of the means and selected quantiles of the distributions of individual equivalent income, after adjusting for inflation and differences in household size and composition. In Panel (b), we display the evolution of social welfare per capita based on the Gini social welfare function.

ranking, but is nonetheless a useful first step.[17] One insight from Table 5.1 is that any inequality averse social planner would conclude that social welfare is higher in 2007 than in the previous years. Another finding is that social welfare remains higher after the crisis as compared to 1994, our first year of observation.

In Table 5.2, we examine whether third-degree upward dominance raises the ranking success rate. We find that the use of this refinement matters little, if anything, for the ability to rank income distributions. By contrast, third-degree downward dominance provides an almost complete ranking of the income distributions. As shown in Table 5.3, this ranking criterion resolves all except one of the comparisons that were ambiguous under second-degree dominance. We can also see that these rankings are statistically significant at conventional levels.

Taken together, the findings in Tables 5.2 and 5.3 point to the importance of whether income differences between poorer individuals are viewed as more or less important for social welfare as compared to income differences between richer individuals. If the social planner is more concerned with income differences in the lower part of distributions, weaker criteria than third degree upward dominance are required to make unambigouous conclusions about the changes in the distribution of income over the business cycle.[18] However, if the social planner focuses attention on income differences in the upper part of the distribution, as in the recent studies of the evolution of top incomes, a nearly complete ranking of income distributions can be achieved.

[17]The high rate of success in ranking income distributions by second-degree dominance contrasts with the findings in some other datasets (see Atkinson, 2008).

[18]Appendix Table A.1 shows the necessary degree of upward and downward dominance to achieve a complete ranking. The results show that the degree of upward dominance has to be quite high to raise the ranking success rate substantially.

Table 5.1: Ranking of income distributions by 2nd-degree dominance

YEAR	1995	1996	1997	1998	1999	2000	2001	2005	2006	2007	2008	2009	2010
1994			$<^{a}$					<	<	<	<	$<^{a}$	<
1995		$<^{a}$	<					<	<	<	<	$<^{a}$	
1996										<			
1997										<			
1998						<		<	<	<	<		
1999						<		<	<	<	<		
2000								<	<	<	$<^{a}$		
2001								<	$<^{a}$	<	$<^{a}$		
2005										$<^{a}$			>
2006										<			>
2007											>	>	>
2008													>
2009													

Note: The table makes pairwise yearly comparisons of the income distributions over the period 1994–2010. We report the succesful rankings based on 2nd-degree dominance. We denote by "<" when the later year dominates the earlier year, and with ">" when the earlier year dominates the later year. ^{a} $p > 0.10$. ^{b} $p > 0.05$.

In particular, it is then clear that social welfare steadily increased until 2007, and that the recession caused a reversion in social welfare to the level of year 2000. This can be seen clearly in Panel (b) of Figure 5.1, which shows the estimated social welfare in each year as evaluated by the least restrictive social welfare function that ranks consistently with third-degree downward dominance (i.e. the Gini social welfare function). At its peak in 2007, the equally distributed equivalent income is above € 24,000 per capita and the welfare loss due to inequality is about € 11,400 per capita.

5.2 Evaluating the distributional effects of policy

To illustrate the usefulness of our framework for policy evaluations, we now apply it to Connecticut's Jobs First experiment.[19] This randomized controlled trial assigned 2,396 welfare recipients to Jobs First, while 4,803 recipients were assigned to Aid for Dependent Children (AFDC). Compared to the high implicit tax rates and no time limit of the AFDC program, Jobs First expanded the earnings disregard and imposed a strict 21-month time limit on welfare participation. Under AFDC, the monthly earnings disregard was $120 in the first year and $90 thereafter, while statutory benefit reduction was 66 % in the first four months, and 100 % thereafter.[20] In contrast, Jobs First entailed no benefit reduction below the federal poverty line and a

[19] For detailed information about the program and for descriptive statistics, we refer to Bloom, Scrivener, Michalopoulos, Morris, Hendra, Adams-Ciardullo, and Walter (2002) or Bitler, Gelbach, and Hoynes (2006).

[20] Due to several expense disregards, lags in enforcement and the implicit wage subsidy from the Earned Income Tax Credit, Bitler, Gelbach, and Hoynes (2006) estimate the effective benefit reduction at about 33 %.

Table 5.2: Ranking of income distributions by 3rd-degree upward dominance

YEAR	1995	1996	1997	1998	1999	2000	2001	2005	2006	2007	2008	2009	2010
1994	<[a]	<[a]	<					<	<	<	<	<	<
1995		<	<					<	<	<	<	<	
1996										<			
1997										<			
1998					>[a]	<		<	<	<	<		
1999						<		<	<	<	<		
2000								<	<	<	<		
2001								<	<	<	<		
2005										<			>
2006										<			>
2007											>	>	>
2008													>
2009													

Note: The table makes pairwise yearly comparisons of the income distributions over the period 1994–2010. We report the succesful rankings based on 3rd-degree upward inverse stochastic dominance. We denote by "<" when the later year dominates the earlier year, and with ">" when the earlier year dominates the later year. [a] $p > 0.10$. [b] $p > 0.05$.

Table 5.3: Ranking of income distributions by 3rd-degree downward dominance

YEAR	1995	1996	1997	1998	1999	2000	2001	2005	2006	2007	2008	2009	2010
1994	>		<	<	<	<	<	<	<	<	<	<	<
1995		<	<	<	<	<	<	<	<	<	<	<	<
1996			<	<	<	<	<	<	<	<	<	<	<
1997				<	<	<	<	<	<	<	<	<	<
1998					<[b]	<	<	<	<	<	<	<	<
1999						<	<	<	<	<	<	<	<
2000							<	<	<	<	<	>	>
2001								<	<	<	<	>	>
2005									>	<	>	>	>
2006										<	>	>	>
2007											>	>	>
2008												>	>
2009													<

Note: The table makes pairwise yearly comparisons of the income distributions over the period 1994–2010. We report the succesful rankings based on 3rd-degree downward inverse stochastic dominance. We denote by "<" when the later year dominates the earlier year, and with ">" when the earlier year dominates the later year. [a] $p > 0.10$. [b] $p > 0.05$.

100 % reduction beyond this.[21]

Bitler, Gelbach, and Hoynes (2006) evaluated how the Jobs First-program affected the distribution of earnings, transfers and total income among participants. In line with the predictions from economic theory, the estimated quantile treatment effects reveal considerable heterogeneity in the impact of the program. To evaluate whether this program was an overall success, we extend on their analysis by using our framework to rank the actual and counterfactual income distributions and to quantify the difference in social welfare between the two distributions.

Table 5.4 displays the results. In panel A, we report the degree of upward and downward dominance necessary to rank the distributions of total income under Jobs First and AFDC.[22] By identifying the least restrictive member of the parametric social welfare functions that rank consistently with the estimated degree of dominance, we can also compute the social welfare level of the dominating distribution as compared to the dominated distribution. Panel B reports the percentage increase in social welfare in the dominating distribution. To ease the interpretation of the social preferences underlying the dominance results, panel C illustrates the weight functions of the least restricive members of the parametric social welfare functions. For brevity, we report the ratios of the weights of the median individual compared to the the 5% poorest, the 30% poorest, the 30% richest, and the 5% richest.

We can see that a refinement of second-degree dominance is necessary to rank the income distributions under Jobs First and AFDC. The first column shows that we need a high degree of upward dominance to reach an unambiguous ranking. If the social planner is sufficiently averse to income differences in the lower part of the distributions, AFDC unambiguously provides higher social welfare than Jobs First. For instance, the least restrictive member of the parametric welfare functions that rank consistently with 9th-degree upward dominance assigns about ten times as much weight to the 30th percentile compared to the median income. With such social preferences, Jobs First is estimated to reduce social welfare by 14.4 percent.

The second column confirms the ability of third-degree downward dominance to resolve comparisons that were ambiguous under second-degree dominance. If the social planner supports the principle of first-degree UPTS, the Jobs First distribution dominates the AFDC distribution. This implies that an unambiguous conclusion can be drawn with quite unrestrictive social preferences; for example, it is sufficient to assign 1.4 times as much weight to the 30 percentile compared to the median income. By applying the least restrictive member of the parametric welfare functions that ranks consistently with third degree downward dominance, we estimate that Jobs First increases social welfare by almost 11 percent.

[21]Compared to AFDC, Jobs First also expanded the work requirement, the asset limit and transitional Medicaid, while enforcing stricter sanctions for violations (see Bloom, Scrivener, Michalopoulos, Morris, Hendra, Adams-Ciardullo, and Walter, 2002).

[22]Total income includes transfers and earnings. Because we do not have access to the micro data, our ranking is based on the estimated quantile treatment effects from Bitler, Gelbach, and Hoynes (2006). As a consequence, we are not able to draw statistical inference about the degree of dominance.

Table 5.4: Comparing income distributions under Jobs First and AFDC

	Upward dominance	**Downward dominance**
A. Ranking		
Degree of dominance	9	3
Dominating distribution	AFDC	Jobs First
B. Social welfare gains		
ΔW_P	14.41%	10.94%
C. Weights at quantiles		
p(.05)	89.39	1.90
p(.30)	10.54	1.40
p(.70)	0.03	0.60
p(.95)	0.00	0.10

Note: The results are based on the estimated quantile treatment effects reported in Bitler, Gelbach, and Hoynes (2006). Income is defined as the sum of earnings and transfers, averaged over quarters 1–16. Panel A shows the degree of upward and downward dominance necessary to rank the income distributions under Jobs First and AFDC. Panel B shows the gains in social welfare (in percent) in the dominating distribution. We report the change in social welfare for the least restrictive member of the parametric social welfare functions that rank consistently with the estimated degree of dominance. Panel C displays the weights at selected quantiles relative to the weight at the median income for the least restrictive member of the parametric social welfare functions.

6 Conclusion

Since the seminal contributions of Kolm (1969) and Atkinson (1970), second-degree dominance has become a widely accepted criterion for ranking distribution functions. But in many applications where the distribution functions intersect, a reasonable refinement of this criterion is necessary to attain an unambiguous ranking. Although the theoretical literature offers dominance criteria of an arbitrary order, they are rarely used; the reason is that higher degree dominance criteria are generally viewed as difficult to interpret and hard to justify because they rely on assumptions about higher order derivatives (see e.g. Atkinson, 2003; Atkinson, 2008). To address these concerns, a large and growing literature has explored third degree dominance as a criterion for ranking distributions.

Our paper contributes by providing a general framework to unambiguously rank any set of distribution functions and quantify the social welfare level of a dominating distribution as compared to a dominated distribution. Our framework is based on two complementary sequences of nested inverse stochastic dominance criteria. The first sequence includes the traditional inverse dominance criteria of third and higher degrees; it is called *upward dominance* because it aggregates the integrated inverse distribution function from below, and therefore places more emphasis on differences that occur in the lower part of the distributions. The second sequence is novel and complements the traditional criteria by placing more emphasis on differences that occur in the upper part of the distribution; we call it *downward dominance* because it aggregates the integrated inverse distribution function from above. Since the sequences are hierarchical, the sensitivity to differences in the lower (upper) part of the distribution increases with the degree of upward (downward) dominance. The two sequences coincide at second-degree dominance, and thus both satisfy the Pigou-Dalton transfer principle.

For each sequence, we show equivalence in the ranking of distributions according to the dominance criteria and a general family of rank-dependent social welfare functions. Because the sequences of dominance criteria are nested, our equivalence results allow us to uniquely identify the largest subfamily of welfare functions – and thus the least restrictive social preferences – that give an unambiguous ranking of any set of distribution functions. We also provide a characterization of the largest subfamily of social welfare functions that rank consistently with dominance of any given degree. Because of the equivalence result, this characterization provides intepretation and justification not only for the social welfare functions, but also for the use of higher degree dominance criteria in comparison of distribution functions. We further show that the members of two alternative parametric families of social welfare functions can be divided into subfamilies according to their relationship with the nested inverse stochastic dominance criteria. The parametric families are well known, easily implementable and the estimated social welfare can be given a money metric interpretation.

We show the usefulness of our framework with two empirical applications. The first uses data from the UK to study how the distribution of household income evolved over a boom and a bust era in the British economy. We show how our framework can be used to make unambiguous statements about the social welfare implications of the changes in the household income distribution over the business cycle. The second uses random-assignment data to evaluate the distributional effects of Connecticut's Jobs First program, which involved generous earnings disregard and strict time limits. We use our framework to infer the least restrictive social preferences that allow an unambiguous conclusion of whether this program was an overall success. In both applications, we find that third-degree downward dominance is a particularly powerful refinement of second-degree dominance, providing an almost complete ranking of the distribution functions. By comparison, the traditional criterion of third-degree upward dominance resolves few of the comparisons that were ambiguous under second-degree dominance.

References

AABERGE, R. (2000): "Characterizations of Lorenz curves and income distributions," *Social Choice and Welfare*, 17, 639–653, 10.1007/s003550000046.

——— (2001): "Axiomatic Characterization of the Gini Coefficient and Lorenz Curve Orderings," *Journal of Economic Theory*, 101(1), 115–132.

——— (2009): "Ranking intersecting Lorenz curves," *Social Choice and Welfare*, 33(2), 235–259.

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.

ATKINSON, A. (1970): "On the measurement of inequality," *Journal of Economic Theory*, 2(3), 244 – 263.

——— (2003): "Multidimensional Deprivation: Contrasting Social Welfare and Counting Approaches," *The Journal of Economic Inequality*, 1(1), 51–65.

——— (2008): "More on the measurement of inequality," *The Journal of Economic Inequality*, 6(3), 277–283.

ATKINSON, A., AND T. PIKETTY (eds.) (2007): *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*. Oxford University Press.

——— (eds.) (2010): *Top Incomes: A Global Perspective*. Oxford University Press.

BARRETT, G. F., AND S. G. DONALD (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.

BITLER, M. P., J. B. GELBACH, AND H. W. HOYNES (2006): "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments," *American Economic Review*, 96(4), 988–1012.

——— (2008): "Distributional impacts of the Self-Sufficiency Project," *Journal of Public Economics*, 92(3-4), 748–765.

BLOOM, D., S. SCRIVENER, C. MICHALOPOULOS, P. MORRIS, R. HENDRA, D. ADAMS-CIARDULLO, AND J. WALTER (2002): *Jobs First: final report on Connecticut's welfare reform initiative : summary report*. Manpower Demonstration Research Corp.

BLUNDELL, R., AND B. ETHERIDGE (2010): "Consumption, Income and Earnings Inequality in Britain," *Review of Economic Dynamics*, 13(1), 76–102.

CHEW, S. H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, 51(4), 1065–92.

CHIU, W. (2007): "Intersecting Lorenz Curves, the Degree of Downside Inequality Aversion, and Tax Reforms," *Social Choice and Welfare*, 28(3), 375–399.

DARDANONI, V., AND P. LAMBERT (1988): "Welfare rankings of income distributions: A role for the variance and some insights for tax reform," *Social Choice and Welfare*, 5, 1–17.

DAVIDSON, R., AND J.-Y. DUCLOS (2000): "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 68(6), 1435–1464.

DAVIES, J., AND M. HOY (1994): "The Normative Significance of Using Third-Degree Stochastic Dominance in Comparing Income Distributions," *Journal of Economic Theory*, 64(2), 520 – 530.

——— (1995): "Making Inequality Comparisons When Lorenz Curves Intersect," *American Economic Review*, 85(4), 980–86.

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.

DONALDSON, D., AND J. A. WEYMARK (1980): "A single-parameter generalization of the Gini indices of inequality," *Journal of Economic Theory*, 22(1), 67–86.

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.

EBERT, U. (1987): "Size and distribution of incomes as determinants of social welfare," *Journal of Economic Theory*, 41(1), 23–33.

FIRPO, S. (2007): "Efficient Semiparametric Estimation of Quantile Treatment Effects," *Econometrica*, 75(1), 259–276.

FIRPO, S., N. M. FORTIN, AND T. LEMIEUX (2009): "Unconditional Quantile Regressions," *Econometrica*, 77, 953–973.

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

——— (1980): "Continua of stochastic dominance relations for unbounded probability distributions," *Journal of Mathematical Economics*, 7(3), 271 – 285.

FISHBURN, P. C., AND R. D. WILLIG (1984): "Transfer principles in income redistribution," *Journal of Public Economics*, 25(3), 323 – 328.

HADAR, J., AND W. R. RUSSELL (1969): "Rules for Ordering Uncertain Prospects," *American Economic Review*, 59(1), 25–34.

HÁJEK, J., AND Z. ŠIDÁK (1967): *Theory of rank tests*, Probability and mathematical statistics. Academic Press.

HARDY, G., J. LITTLEWOOD, AND G. PÓLYA (1934): *Inequalities*, Cambridge Mathematical Library. Cambridge University Press.

IMBENS, G. W., AND W. K. NEWEY (2009): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 77(5), 1481–1512.

KOLM, S. (1969): "The optimal production of social justice," in *Public Economics*. Julius Margolis and Henri Guitton.

——— (1976): "Unequal inequalities. I," *Journal of Economic Theory*, 12(3), 416–442.

LAMBERT, P. (1993): *The distribution and redistribution of income: a mathematical analysis*. Manchester University Press.

LE BRETON, M., A. MICHELANGELI, AND E. PELUSO (2012): "A stochastic dominance approach to the measurement of discrimination," *Journal of Economic Theory*, 147(4), 1342–1350.

LE BRETON, M., AND E. PELUSO (2009): "Third-degree stochastic dominance and inequality measurement," *Journal of Economic Inequality*, 7(3), 249–268.

LINTON, O., E. MAASOUMI, AND Y.-J. WHANG (2005): "Consistent Testing for Stochastic Dominance under General Sampling Schemes," *Review of Economic Studies*, 72(3), 735–765.

MEHRAN, F. (1976): "Linear Measures of Income Inequality," *Econometrica*, 44(4), pp. 805–809.

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

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

SEN, A. (1974): "Informational bases of alternative welfare approaches : Aggregation and income distribution," *Journal of Public Economics*, 3(4), 387–403.

SHORROCKS, A. F. (1983): "Ranking Income Distributions," *Economica*, 50(197), 3–17.

SHORROCKS, A. F., AND J. E. FOSTER (1987): “Transfer Sensitive Inequality Measures,” *Review of Economic Studies*, 54(3), 485–97.

WHITMORE, G. A. (1970): “Third-Degree Stochastic Dominance,” *American Economic Review*, 60(3), 457–59.

YAARI, M. E. (1987): “The Dual Theory of Choice under Risk,” *Econometrica*, 55(1), 95–115.

——— (1988): “A controversial proposal concerning inequality measurement,” *Journal of Economic Theory*, 44(2), 381 – 397.

ZOLI, C. (1999): “Intersecting generalized Lorenz curves and the Gini index,” *Social Choice and Welfare*, 16(2), 183–196.

——— (2002): “Inverse stochastic dominance, inequality measurement and Gini indices,” *Journal of Economics*, 9(1), 119–161.

A Appendix: Proofs and supplementary tables

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 of Theorem 2.1.[23] 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

[23] 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.

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

$$
\begin{aligned}
H_1(v,h_1) &= P'(v) - P'(v+h_1) && \text{(A.2)}\\
H_2(s,t,h_1) &= H_1(s,h_1) - H_1(t,h_1) && \text{(A.3)}\\
H_{j+1}\left(s,t,h_1,h_2,\ldots,h_j\right) &= H_j\left(s,t,h_1,h_2,\ldots,h_{j-1}\right) && \text{(A.4)}\\
&\quad -H_j\left(s+h_j,t+h_j,h_1,h_2,\ldots,h_{j-1}\right) \quad \text{for } j=2,3,\ldots && \text{(A.5)}
\end{aligned}
$$

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

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

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 (2.1) 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 of Theorem 2.4. The proof follows exactly the reasoning used in the proof of Theorem 2.3, using the following expression,

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

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

Proof of Theorem 4.2. Let

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

and note that

$$2\sum_{j=1}^{N}\frac{\sin(j\pi s)\sin(j\pi t)}{(j\pi)^2}=s(1-t) \tag{A.10}$$

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

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

where

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

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

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

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

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

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

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

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

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

□

Proof of Corollary 4.2. Theorem 4.1 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 (4.4), it follows from Theorem 4.2 that the first term $(1-u)^{i-2}\int_0^1 Q_n(t)\, dt$ of expression (4.3) converges to a process that has the same distribution as $\sum_{j=1}^{\infty}\frac{\sqrt{2}}{j\pi}\left[(1-u)^{i-2}\int_0^1\frac{\sin(j\pi t)}{f\left(F^{-1}(t)\right)}dt\right]Z_j$, while the second term $\left[\int_u^1(t-u)^{i-2}Q_n(t)\,dt\right]$ of expression (4.3) converges to a process that has the same distribution as $\sum_{j=1}^{\infty}\frac{\sqrt{2}}{j\pi}\left[\int_u^1(t-u)^{i-2}\frac{\sin(j\pi t)}{f\left(F^{-1}(t)\right)}dt\right]Z_j$. □

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.