Unambiguous Comparison of Intersecting Distribution Functions*

Rolf Aaberge† Tarjei Havnes‡ Magne Mogstad§

January 4, 2012

PRELIMINARY VERSION: DO NOT CITE

Keywords:

JEL-codes:

*The project is part of the research activities at the ESOP center at the Department of Economics, University of Oslo. ESOP is supported by The Norwegian Research Council.

†Statistics Norway

‡Univ of Oslo, ESOP & Statistics Norway

§University College London, Statistics Norway & ESOP

1 Introduction

How do we compare intersecting distribution functions? This question is important for both policy, descriptive analysis and causal inference. First, a key task of statistical offices and government agencies is to compare distribution functions across countries, subgroups and time. Second, descriptive research devotes much attention to the comparison of distributions of economic variables, such as wages, income, consumption and wealth, as they are viewed as important determinants of economic welfare as well as markers for what kind of activities that are rewarded in an economy. Third, there is a growing interest in empirical research in how to estimate not only the mean impact of a policy change, but also the counterfactual outcome distribution in the absence of a policy intervention.[1] At the same time, little attention has been devoted to how to use the additional information, that is, how to compare the counterfactual and actual outcome distribution in a way that allows researchers and policymakers to move beyond differences in means when assessing the gains and losses of a policy.

This paper proposes a general framework for comparing distribution functions, in situations where they intersect and no unambiguous ranking can be attained without introducing weaker ranking criteria than second degree stochastic dominance. Although the theoretical literature offers dominance criteria of third degree and higher to rank intersecting distribution functions,[2] they are generally viewed as difficult to interpret and hard to justify because they depend on assumptions about third and higher order derivatives (see e.g. Atkinson, 2003). Thus, most empirical studies rely exclusively on one or a few social welfare functions to compare the mean and inequality of intersecting distribution functions. A natural concern is that the conclusions reached are sensitive to the more or less arbitrary choice of social welfare function. It is, therefore, due time to start bridging the wide gap between the theoretical and the empirical strand of the literature concerned with the comparison of intersecting distribution functions. That is the focus of this paper. We now describe step-by-step the content and the contributions of our paper.

POINT OF DEPARTURE. We begin by clarifying the relationship between second degree stochastic dominance and social welfare functions in the ranking of distribution distribution functions. As is well known, all inequality averse social welfare functions rank distribution functions consistently with second degree stochastic dominance. But in many applications,

[1] A number of recent studies have demonstrated that uncovering the counterfactual outcome distributions can be useful to assess the effects of policy interventions when theory predicts heterogenous responses, by estimating quantile treatment effects (see e.g. Bitler et al. 2006; 2008; Firpo (2007); Firpo, Fortin, and Lemieux (2009)), marginal treatment effects Heckman and Vytlacil (2005); Carneiro, Heckman, and Vytlacil (2010), decomposition methods (see e.g. DiNardo, Fortin, and Lemieux, 1996; Juhn, Murphy, and Pierce, 1993; Fortin, Lemieux, and Firpo, 2011) and structural models (see e.g. Blundell, Brewer, Haan, and Shephard, 2009; Aaberge and Colombino, 2010). A common feature of these methods is that they move beyond mean impacts, examining whether the effect of a policy change is constant across the distribution, or whether it leads to larger changes in certain parts of the distribution.

[2] See e.g. Anderson (1996), Barrett and Donald (2003), and Davidson and Duclos (2000) for a discussion of higher order stochastic dominance criteria.

weaker dominance criteria are required to attain an unambigous ranking of distribution functions. To deal with such situations, we employ two alternative sequences of nested inverse stochastic dominance criteria, and characterize their relationship to a general family of social welfare functions in the ranking of distribution functions.

SEQUENCES OF DOMINANCE CRITERIA. The first sequence includes the traditional inverse dominance criteria of third and higher degrees; it is called *upward dominance* because it places more emphasis on differences that occur in the lower part of the distributions.[3] The second sequence is novel and complements the traditional criteria in placing more emphasis on differences that occur in the upper part of the distribution; we therefore call it *downward dominance.* Because 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.

ASYMPTOTICS AND INFERENCE. To make statistical inference about upward and downward dominance of any degree, we develop distribution theory for tests based on the empirical distribution functions where it is demonstrated that the associated empirical process converges in distribution to a Gaussian process. Thus, the empirical dominnance criteria are asymptotically normally distributed both when considered as a process and for fixed rank in the distribution.

INTERPRETATION OF DOMINANCE CRITERIA. 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 can be interpreted as employing the Gini social welfare function to compare social welfare of those located at the lower tail of each quantile of the distributions.[4] As a consequence, we do not have rely on assumptions about third and higher degree derivatives to interpret the sequences of dominance criteria.

EQUIVALENCE RESULT. We next characterize the relationship between upward and downward dominance and social welfare functions in the ranking of distribution 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 (1988; 1987), 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 the social planner who employs the family of social welfare functions to compare intersecting distribution functions.

Because the sequence of dominance criteria are nested, our equivalence result allows us to uniquely identify the largest subfamily of welfare functions - and thus the least restrictive social preferences - required to reach an unambiguous ranking of any set of distribution functions. In

[3] See Muliere and Scarsini (1989).

[4] The Gini social welfare function was originally introduced by Sen (1974).

doing so, we bridge the dominance approach and the social welfare function approach to the comparison of intersecting distribution functions.

NORMATIVE JUSTIFICATION. We also provide a complete axiomatic characterization of the largest subfamily of social welfare functions that rank consistently with dominance of any given degree. Because of the equivalence result, the characterization gives a normative justification not only for the social welfare functions, but also for the use of higher degree dominance criteria in comparison of distribution functions. The subfamily associated with upward dominance is characterized by (generalizations of) the principle of downside positional transfer sensitivity (see Zoli (1999) and 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.

PARAMETRIC SUBFAMILIES. To not only answer whether one distribution is better than another distribution, but also get an estimate of by how much, we extend our framework. To this end, it is necessary to work with parametic social welfare functions We show that the members of two different parametric families of social welfare functions can be divided into subfamilies according to their relationship to 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 determine the distribution function, no information is lost by restricting the focus to the parametric social welfare functions.

EMPIRICAL APPLICATIONS. To demonstrate the usefulness of this framework for empirical research, we consider two applications. The first application uses random-assignment data to compare the costs and benefits of Connecticut's Jobs First program, which involved generous earnings disregard and strict time limits. 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, and reduced income for a sizable group after the time-limits took effect. To evaluate whether this program was an overall success, we use our framework to aggregate the gains and losses and compare them to the financial costs (including cash transfers, administrative costs and operating costs) in a coherent way.

The second application uses data from Norway and the United States to study how the distribution of earnings has evolved over the last few decades. The evolution of the earnings distribution is one of he most extensively researched topics in economics. In a widely cited review of the literature, Gottschalk and Smeeding (1997) conclude that the earnings distribution has become more dispersed in most OECD countries during the 1980s and early 1990s. Moreover, they argue that many countries with fairly low levels of inequality experienced some of the largest increases in inequality. These conclusions rest on numerous empirical studies relying exclusively on one or a few summary measures. A concern is, however, that the conclusions reached are sensitive to the more or less arbitrary choice of summary measure. We show how our framework can be used to make unambiguous statements that are economically interpretable interpretable about the evoluation of the earnings distribution over time and between countries.

OUTLINE. Section 2 characterizes the relationship between inverse stochastic dominance criteria and social welfare functions in the ranking of distribution distribution functions. Section 3 identifies and describes the parametric families that rank consistent with upward and downward dominance. Section 4 presents the distribution theory. Section 5 presents the empirical analysis. Section 6 summarizes the main results of the paper and relates them to alternative approaches to comparing intersecting distribution functions.

2 Inverse stochastic dominance and rank-dependent social welfare

This section begins by presenting the general family of social welfare functions. We next characterize the relationship between third degree upward and downward dominance and social welfare functions in the ranking of distribution functions. Lastly, we introduce the full hierarchical sequences of nested inverse stochastic dominance criteria, and show how it allows us to uniquely identify the largest subfamily of welfare functions required to reach an unambiguous ranking of any set of distribution functions.

2.1 Rank-dependent social 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 functions.

The basic axioms

Assume that the preference relation $\succeq$ of the social planner satisfies the following axioms:

Axiom 1. ***(Order)***. *$\succeq$ is a transitive and complete ordering on $\mathscr{F}$.*

Axiom 2. ***(Continuity)***. *For each $F \in \mathscr{F}$ the sets $\{F^\star \in \mathscr{F} : F \succeq F^\star\}$and $\{F^\star \in \mathscr{F} : F^\star \succeq F\}$ are closed (w.r.t. L_1-norm).*

Axiom 3. ***(Dominance)***. *Let $F_0, F_1 \in \mathscr{F}$. If $F_1^{-1}(t) \geq F_0^{-1}(y)$ for all $t \in [0,1]$ and the inequality holds strictly for some t$\in (0,1)$ then $F_1 \succeq F_0$.*

Axiom 4. ***(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 $\left(\alpha F_1^{-1} + (1-\alpha)F_2^{-1}\right)^{-1} \succeq \left(\alpha F_0^{-1} + (1-\alpha)F_2^{-1}\right)^{-1}$.*

The first three axioms are standard ordering, continuity and dominance assumptions. Axiom 4 is analogous to the independence axiom underlying Atkinson's (1970) expected utility type of social welfare functions, which requires that the ordering of distributions of income is invariant with respect to identical mixing of the *distributions* being compared; that is, mixing of population shares given income levels. The dual independence axiom requires instead that the ordering 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. For further discussion, see Yaari (1988) and Aaberge (2001).

The general family

Yaari (1988; 1987) proved that the preference relation $\succeq$ that satisfies Axioms 1–4 can be represented by the following 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 that is a member of the following set of preference functions

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

As demonstrated by Yaari (1988), second degree inverse stochastic dominance is characterized by the social welfare functions W_P if and only if $P'(t) > 0$ and $P''(t) < 0$.

Definition 2.1. A distribution function F_1 is said to *second* degree *stochastic 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 $u \in (0,1)$.

If $\mu_{F_1} = \mu_{F_0}$, the condition of second degree inverse 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.[5]

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.

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 $u = .6$. Note that the preference function must be concave and lie above the diagonal to ensure that W_P is inequality averse and satisfy the Pigou-Dalton transfer principle.

Normative justification

The normative justification 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 an 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:

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

Thus, as was recognized by 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 the 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 distribution 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.

[5] As was demonstrated by Hardy, Littlewood, and Pólya (1934), Kolm (1969) and Atkinson (1970), second degree inverse stochastic is equivalent to second degree stochastic dominance. Moreover, under the restriction of equal mean incomes second degree inverse stochastic dominance is equivalent to the criterion of non-intersecting Lorenz curves.

Parametric subfamilies

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 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. Inserting for (2.5) in (2.3), W_P becomes equal to the extended Gini family of social welfare functions defined by

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

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 for (2.8) in (2.3), W_P becomes equal to the extended Lorenz family of social welfare functions

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

Since $\left\{\mu_F, W_{G_k}(F) : k = 3,4,...\right\}$ 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.

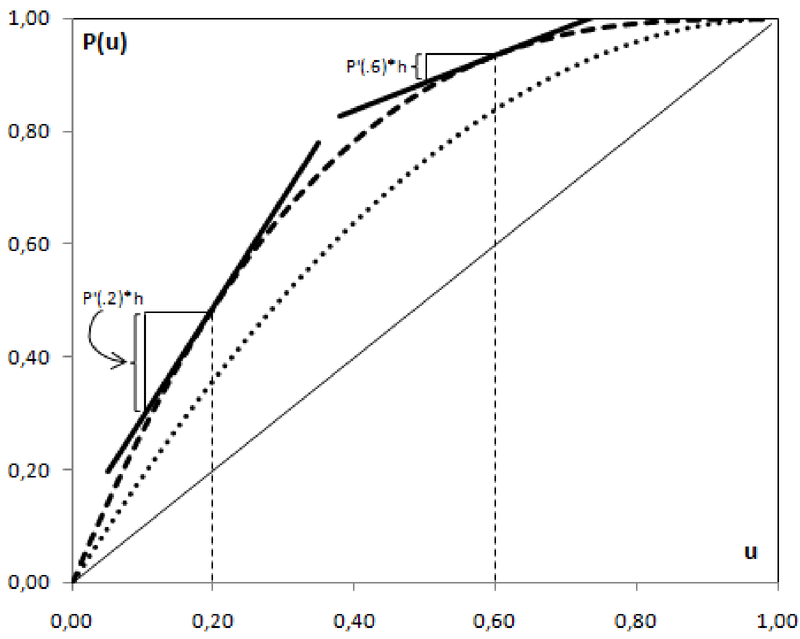


Figure 2.1: Examples of the preference function $P(\cdot)$ that preserves 3rd (dotted) and 4th degree (dashed) upward inverse stochastic dominance.

Note: The weight assigned to individuals at rank u equals the derivative of P at u.

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

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

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 from inserting (2.10) for Λ_F^2 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

$$\Lambda^3_{F_1}(u) \geq \Lambda^3_{F_0}(u) \text{ for all } u \in [0,1]$$

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

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

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 uniquely determined by the Gini social welfare function, which means that application of the criterion of third-degree upwards inverse stochastic dominance corresponds to compare the social welfare of the individuals located at the lower tail of each quantile of the distributions.

Proposition 2.1. *Let F_1 and F_0 be members of 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

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

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

Transfer principle

To provide a normative justification for dominance 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. However, an alternative version of the principle of diminishing transfers introduced by Mehran (1976) – and called the principle of positional transfer sensitivity by Zoli (1999) – proves 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 $\Delta^1_{st} W_P(\delta,h) \equiv \Delta_s W_P(\delta,h) - \Delta_t W_P(\delta,h)$. We can then define the principle of first degree downside position transfer sensitivity.

Definition 2.3. W_P satisfies the principle of first degree downside positional transfer sensitivity (DPTS) if and only if

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

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. This is similar in spirit to the principle of diminishing transfers. However, because income distributions are usually skewed to the right, a fixed difference in ranks implies that transfers in the upper part of the income distribution tend to take place between individuals with larger differences in income, as illustrated in Figure 2.2. An inequality averse social planner who supports the principle of first degree *DPTS* is said to exhibit downside positional inequality aversion of first degree.

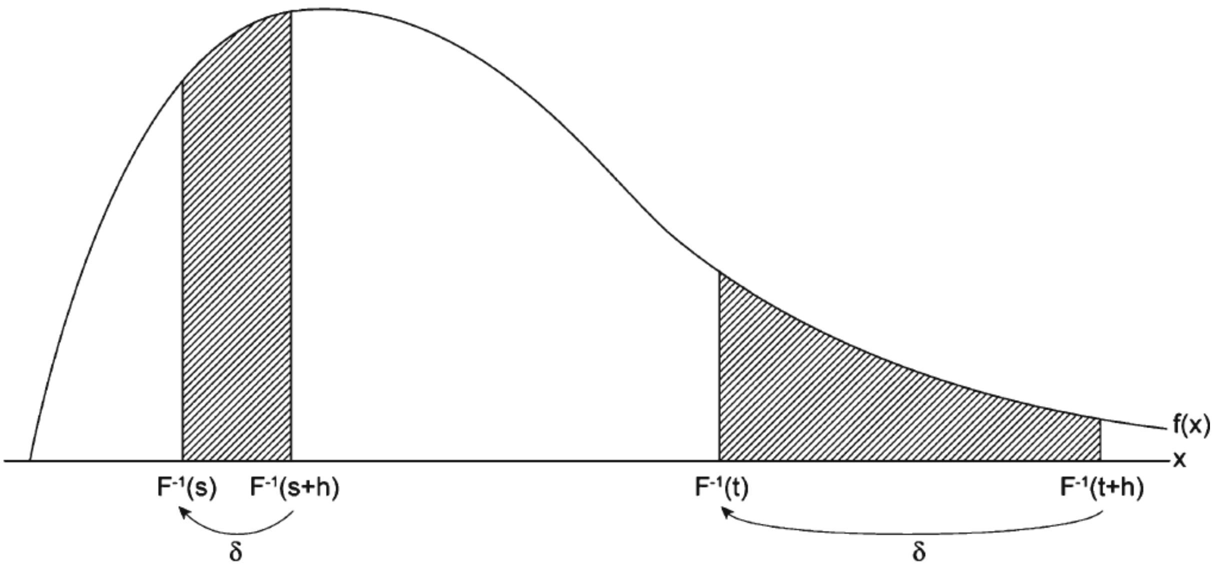


Figure 2.2: Downside positional transfer sensitivity

Equivalence result

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

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

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 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.[6]

[6]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 (i) and (iii) in Theorem 2.1. Aaberge (2000) demonstrates that J_P defined by (2.2) satisfies the principle of diminishing transfers under conditions that depend on the shape of the preference function P *as well as* the shape of the income distribution F.

2.2.2 Downward dominance and social welfare

Section 2.3 demonstrated that a social planner who supports the criterion of third degree upward inverse stochastic dominance exhibits aversion to downside inequality. However, in some cases, the social planner might be more concerned with differences in the upper part of the distribution. The recent focus on the evolution of top incomes is one example (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 inverses of the distribution functions 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 in 2.16 and 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

$$\tilde{\Lambda}_{F_1}^3(u) \geq \tilde{\Lambda}_{F_0}^3(u) \text{ for all } u \in [0,1]$$

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

From equation (2.16), it is clear that the criterion of third degree upward 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 K 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

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

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

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

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 uniquely determined by employing the Gini social welfare function to compare the social welfare of the individuals located in the upper tail of each quantile of the distributions.

Proposition 2.2. *Let F_1 and F_0 be members of F. Then the following statements are equivalent:*
(i) F_1 third degree downward inverse stochastic dominates F_0
(ii) $\tilde{\mu}_{F_1}(u)\,(1 - D_3(u;F_1)) \geq \tilde{\mu}_{F_2}(u)\,(1 - D_3(u;F_0))$ 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

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

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

Propositions 2.1 and 2.2 highlight the similarities and differences between upwards and downwards inverse stochastic dominance: The criterion of upward dominance is a sequential comparison of the mean and inequality among the poorest u percent of the population, whereas the criterion of downward dominance is a sequential comparison of the mean and inequality among the richest $(1-u)$ percent of the population.

Transfer principle

To provide a normative justification for 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 analysing 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. Further, let $\Delta^1_{st} W_P(\delta,h) \equiv \Delta_s W_P(\delta,h) - \Delta_t W_P(\delta,h)$. We can then define the principle of first degree upside position transfer sensitivity.

Definition 2.5. W_P satisfies the principle of first degree upside positional transfer sensitivity (UPTS) if and only if

$$\Delta^1_{st} W_P(\delta,h) < 0, \quad \text{when } s < t.$$

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.

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

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

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 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 Theorem 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 many cases, neither upward nor downward dominance of third degree allows an unambiguous ranking of the distribution functions under comparison (see for instance the application to *Jobs*

First below). This subsection therefore introduces full hierarchical sequences of nested inverse stochastic dominance criteria that allow 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

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

Definition 2.6. A distribution F_1 is said to ith degree upward inverse stochastic dominate F_0 if and only if

$$\Lambda_{F_1}^i(u) \geq \Lambda_{F_0}^i(u) \text{ for all } u \in [0,1]$$

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

$$\tilde{\Lambda}_{F_1}^i(u) \geq \tilde{\Lambda}_{F_0}^i(u) \text{ for all } u \in [0,1]$$

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. However, as will be demonstrated below the choice between higher degree og upward and downward domincance 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

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 upward and downward dominance of degree i can be given analogous interpretations. 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).

The quantile-specific lower tail 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 quantile-specific lower tail Gini family of social welfare functions can then be expressed as

$$\mu_F(u)\left(1 - G_i(u;F)\right).$$

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 quantile-specific upper tail Lorenz family of social welfare functions can then be expressed as

$$\tilde{\mu}_F(u)\left(1 - D_i(u;F)\right).$$

The following propositions shows that the criterion of i^{th} degree upward (downward) dominance is uniquely determined by employing the Gini (Lorenz) social welfare function of order $(i-1)$ to compare social welfare among individuals located at the lower (upper) tail of each quantile of the distributions.

Proposition 2.3. *Let F_0 and F_1 be members of $\mathscr{F}$. Then the following statements are equivalent:*

(i) F_1 *i^{th} degree upward inverse stochastic dominates F_0*

(ii) $\mu_{F_1}(u)\left(1 - G_i(u;F_1)\right) \geq \mu_{F_0}(u)\left(1 - G_i(u;F_0)\right)$ 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

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

which follows by inserting 2.12 and (2.23) in (2.25). □

Proposition 2.4. *Let $F_0$and F_1 be members of $\mathscr{F}$. Then the following statements are equivalent:*

(i) F_1 i^{th}*degree downward inverse stochastic dominates* F_0

(ii) $\tilde{\mu}_{F_1}(u)\,(1-D_i(u;F_1)) \geq \tilde{\mu}_{F_0}(u)\,(1-D_i(u;F_0))$ 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

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

which follows by inserting (2.17) and (2.24) in (2.26). □

Propositions 2.3 and 2.4 highlight the similarities and differences between upwards and downwards inverse stochastic dominance of degree i: The criterion of upward dominance is a sequential comparison of the mean and inequality among the poorest u percent of the population, while the criterion of downward dominance is a sequential comparison of the mean and inequality among the richest $(1-u)$ percent of the population.

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 upside positional transfer sensitivity (UPTS) of degree i if and only if, for all $k = 1,2,\ldots,i$

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

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

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

According to ith degree UPTS (DPTS), given that two alternative sequences of fixed transfers take place between people with equal difference in ranks, 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.[7] Since UPTS (DPTS) of degree i are stronger criteria than UPTS (DPTS) of degree $i-1$, it seems natural that a social planner that supports the latter will also support the former.

Equivalence result

Let $P^{(i)}$ denote the ith degree derivative of P.

The family of preference functions $\mathscr{P}_i$ is defined by

$$\begin{aligned} \mathscr{P}_i \quad = \quad & \Big\{P \in \mathscr{P} : (-1)^{i-1}P^{(i)}(t) > 0 \text{ with } P^{(j)} \text{ continuous on } (0,1) \\ & \text{and } (-1)^{i-1}P^{(j)}(1) \geq 0 \text{ for all } j = 3,4,\ldots,i-1\Big\} \end{aligned} \tag{2.28}$$

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

$$\begin{aligned} \tilde{\mathscr{P}}_i \quad = \quad & \Big\{P \in \mathscr{P} : P^{(i)}(t) < 0 \text{ with } P^{(j)} \text{ continuous on } (0,1) \\ & \text{and } P^{(j)}(0) \leq 0 \text{ for all } j = 3,4,\ldots,i-1\Big\} \end{aligned} \tag{2.29}$$

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 $i-2$

[7]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.

Proof. In the appendix. □

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

Upward dominance of degree i is ensured by imposing positive (negative) ith degree derivate 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 $(i-1)$ degree upward dominance.

Downward dominance of degree i is ensured by imposing negative ith degree derivate 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 the theorems provides normative justification for ranking distribution functions according to ith degree upward and downward dominance. By comparing (iii) in Theorems 2.3 and 2.4, 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 2.1. Assume (i) F_1 ith degree upward (downward) inverse stochastic dominates F_2 and (ii) F_2 $(i-k)$th degree upward (downward) inverse stochastic dominates F_3. It is immediately evident from Definitions 2.6 and 2.7 that the dominance relations are transitive, in the sense that for $k=0$ (i) and (ii) imply that F_1 ith degree upward (downward) inverse stochastic dominates F_3. Further, from Equations (2.21) and (2.22), it is evident 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. Therefore, for $k=0,1,\ldots$, (i) and (ii) 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 unambigously 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 goes to infinity, 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. Hence, at the limit upward and downward inverse stochastic dominance 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 answer whether one distribution is better than another distribution, but also get an estimate of by how much. 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

Proposition 3.1 arranges the members of the extended 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 consistent with upward dominance of a given degree, and quantify the social welfare level of the dominating distribution as compared to the dominated distribution.

Proposition 3.1. *Let F_1 and F_0 be members of F. Then for $i = 3, 4\dots$*
(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 $k > i$

Proof. Differentiating $P_{1k}(t)$ defined by (2.4) yields

$$P_{1k}^{(j)}(t) = \begin{cases} (-1)^{j-1}\dfrac{k-1!}{(k-j-1)!}(1-t)^{k-j-1}, & j = 1,2,...,k-1 \\ 0 & j = k,k+1,... \end{cases} \tag{3.1}$$

which demonstrates that the derivatives of P_{1k} alternate in sign up to the k-th derivative, $P'_{1k}(t) > 0$ for all $t \in (0,1)$ and $P'_{1k}(1) = 0$. Accordingly, the preference functions $P_{1k}(t)$ for $k = 3,4,...$ are members of $\mathscr{P}_k$ defined by equation (2.28) and thus Proposition 3.1 follows directly from Theorem 2.3. □

Remark. The extended Gini family of social welfare functions has the following properties,
(i) W_{G_i} preserves upward inverse stochastic dominance of degree i and all degrees lower than i, for $i = 3,4,\dots$.
(ii) W_{G_i} obeys the Pigou-Dalton principle of transfers for $i > 2$.
(iii) W_{G_i} obeys the principles of DPTS up to and including $(i-2)$th-degree for $i = 3,4,\dots$.
(iv) The sequence $\{W_{G_i}\}$approaches μ_F as $i \to 2$
(v) The sequence $\{W_{G_i}\}$ approaches the Rawlsian maxi-min criterion as $i \to \infty$.

The left panel of Figure 3.1 displays the preference function $P_{1k}(t)$ when $k = 3$, $k = 4$ and $k = 10$. As we increase the degree of upward dominance preserved by W_{G_i}, 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_i} equals the mean income.

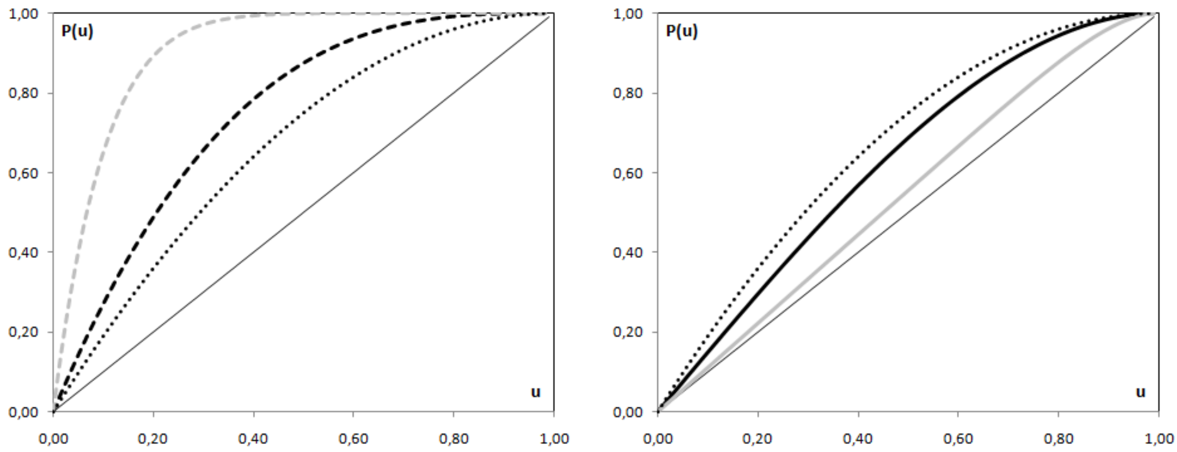


Figure 3.1: Examples of the preference function P that preserves 2nd, 3rd and 10th degree inverse stochastic dominance, upwards (left panel) and downwards (right panel).

Note: The weight assigned to individuals at rank u equal the derivative of P at u. The parametric forms of P is defined in Section 3.

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

Table 3.1: Implied weights under W_{G_i} and W_{D_i}at selected quantiles relative to the weight at the median.

Note: The parametric forms of the weighting function P is defined in Section 3.

Downward dominance and the extended Lorenz family

Proposition 3.2 arranges the members of the extended 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 consistent with downwards dominance of a given degree, and quantify the social welfare level of the dominating distribution as compared to the dominated distribution.

Proposition 3.2. *Let F_1 and F_0 be members of F. Then for $i = 3,4,\dots$*
(i) F_1 ith degree downward inverse stochastic dominates F_0
implies
(ii) $W_{D_k}(F_1) > W_{D_k}(F_0)$ for $k > i$

Proof. Differentiating $P_{2k}(t)$ defined by 2.7 yields

$$P_{2k}^{(j)}(t) = \begin{cases} \dfrac{k-1}{(k-2)}(1-t)^{k-2}, & j=1 \\ -\frac{(k-1)(k-3)!}{(k-j-1)!}t^{k-j-1} & j=2,3,...,k-1 \\ 0 & j=k,k+1,... \end{cases} \tag{3.2}$$

which demonstrates that the higher order derivatives of P_{2k} are negative up to the k-th derivative, $P'_{2k}(t) > 0$ for all $t \in (0,1)$ and $P'_{2k}(1) = 0$. Accordingly, the preference functions $P_{2k}(t)$ for $k = 3,4,...$ are members of $\widetilde{\mathscr{P}}_k$ defined by (2.29) and thus Proposition 3.2 follows directly from Theorem 2.4. □

Remark. The extended Lorenz family of social welfare functions has the following properties,
(i) W_{D_i} preserves downward inverse stochastic dominance of degree i and all degrees lower than i, for $i = 3,4,\dots$.
(ii) W_{D_i} obeys the Pigou-Dalton principle of transfers for $i > 2$.
(iii) W_{D_i} obeys the principles of UPTS up to and including $(i-2)$th-degree for $i = 3,4,\dots$.
(iv) The sequence $\{W_{D_i}\}$ approaches μ_F as $i \to \infty$

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 in the limit as $k \to \infty$. The highest degree of upside inequality aversion occurs as $k \to \infty$, which corresponds to the utilitarian criterion.

Equivalence results

Equivalence with singly/doubly intersecting distribution functions.

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.

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

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

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 $\tilde{\Lambda}^i_F(u)$. Let the empirical process $P_n(u)$ 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

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

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 $0 \le w(u,t) \le 1$ for all $u,t \in [0,1]$ and let $a(u)$ and $b(u)$ be functions of u such that $0 \le a(u) < b(u) \le 1$. 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 $0 \le s \le t \le 1$.

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*

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

Proof. It follows directly from Theorem 4.1 of Doksum (1974) that the empirical process $P_n(t)$ 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), we find 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 for $a(u) = 0$, $b(u) = u$ and $w(u,t) = (u-t)^{i-2} / (i-2)!$ in expression (4.9).

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 processes*

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

which has the same probability distribution as the Gaussian process

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

where $h_j(u)$ *is given by*

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

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 processes*

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

which has the same probability distribution as the Gaussian process

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

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

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

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 (A.10) 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*

$$\begin{aligned} v^2(u,v) &= \sum_{j=1}^{\infty} h_j(u) h_j(v) \qquad (4.5) \\ &= \frac{1}{[(i-2)!]^2} \left\{ 2\int_0^{F^{-1}(u)} \int_0^{y} [(u-F[x])(v-F(y))]^{i-2} F(x)(1-F(y))\,dxdy \right. \\ &\qquad \left. + \int_0^{F^{-1}(v)} \int_0^{F^{-1}(u)} [(u-F[x])(v-F(y))]^{i-2} F(x)(1-F(y))\,dxdy \right\} \end{aligned}$$

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

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

and

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

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 (A.10).

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

$$\begin{aligned} \eta^2(u,v) &= \sum_{j=1}^{\infty} \tilde{h}_j(u) \tilde{h}_j(v) \qquad (4.6) \\ &= [(1-u)(1-v)]^{i-2} \lambda_{22}(0,0) - (1-u)^{i-2} [\lambda_{i2}(v,v) + \gamma_{i2}(0,v)] \\ &\quad - (1-v)^{i-2} [\lambda_{i2}(u,u) + \gamma_{i2}(0,u)] + [\lambda_{ii}(u,v) + \gamma_{ii}(u,v)] \end{aligned}$$

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

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

tends to the $N(0,1)$-distribution for fixed u, where $v^2(u,u)$ is given by (4.5), and the distribution of

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

tends to the $N(0,1)$-distribution for fixed u, where $\eta^2(u,u)$ is given by (4.6).

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

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

where ψ is a continuous nonnegative weight function. By applying Theorems 4.1 and 4.2 and Billingsley (1968, Th. 5.1), we find that K_n converges in distribution to

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

We use the following notation.

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

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

5 Empirical applications

5.1 Quantile treatment effects of the Jobs First program

To illustrate how our framework can be applied in policy evaluations, we now apply it to the US Jobs First program, a randomized trial from the counties of New Haven and Manchester in Connecticut, USA.[8] Between April 1996 and February 1997, all ongoing or open cases in the welfare offices of the two counties were included, for a total available sample of 4 803.

[8]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).

Out of these, 2 396 were assigned to Jobs First, while the remaining were assigned to Aid for Dependent Children (AFDC). Data on monthly welfare and food stamps are available from case files, while quarterly earnings are collected from unemployment insurance files. In addition, data on age, education, race and ethnicity, marital status, and work history is collected in a pre-assignment interview.

Compared to the high implicit tax rates and no time limit of the AFDC program, Jobs First expanded the earnings disregard and enforced 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.[9] This is compared to Jobs First where participants faced no benefit reduction below the federal poverty line and a 100% reduction beyond this.[10]

The effect on work incentives is illustrated in Figure (5.1), which is reproduced from Bitler, Gelbach, and Hoynes (2006) and draws a stylized budget constraint for participants of AFDC and Jobs First, respectively. While AFDC participants faced severe benefit reduction on segment A–B, Jobs First participants faced complete earnings disregard up to the federal poverty line (FPL) on segment A–F before the time limit. Compared to AFDC, Jobs First should therefore be expected to cause significant heterogeneity between individuals: While participants in point A experience a pure substitution effect in favor of working, participants in point C also experience an income effect that may push in the opposite direction. Meanwhile, in point D, participants experience a pure income effect, which decreases labor supply if leisure is a normal good. Further, in points E and H, participants experience a pure substitution effect in favor of leisure. Finally, the expansion of the work requirement implies that some who were able to meet the AFDC work requirement may not be able to meet the work requirement under Jobs First. The loss of transfers will therefore be expected to have a negative effect on total income of these individuals.

In their seminal article, 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 effects reported by Bitler et al. document large heterogeneity in the effects of the program. However, they do not resolve the difficult cases in which effects are non-monotonic and do not satisfy second-degree dominance. As we have discussed, our final evaluation of the welfare consequences of the Jobs first-program then depends crucially on the social welfare weights we associate with the different points in the distribution.

In Panel A of Figure 5.2, we draw the distribution of average quarterly income in quarters 1–16, including earnings and transfers, following treatment assignment among treated and

[9]Due to several expense disregards, lags in enforcement and the implicit wage subsidy from the Earned Income Tax Credit, Bitler et al. (2003) estimate the effective benefit reduction at about 33%.

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

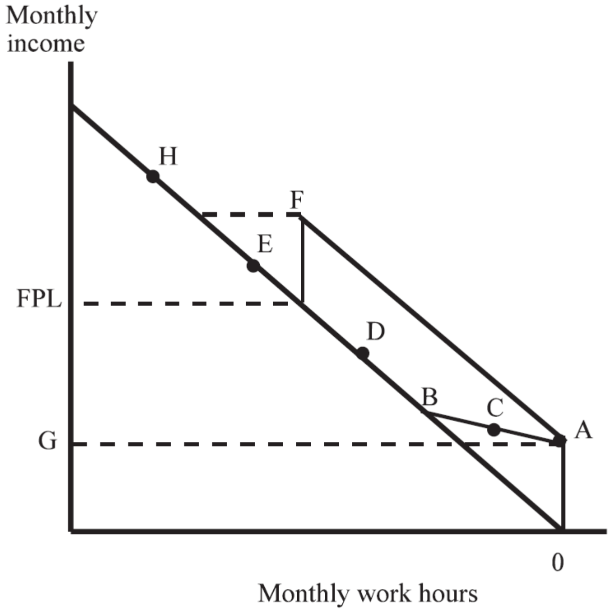


Figure 5.1: Budget constraint under AFDC (segment A–B) and Jobs First (segment A–F).
Note: Reproduced from Bitler, Gelbach, and Hoynes (2006).

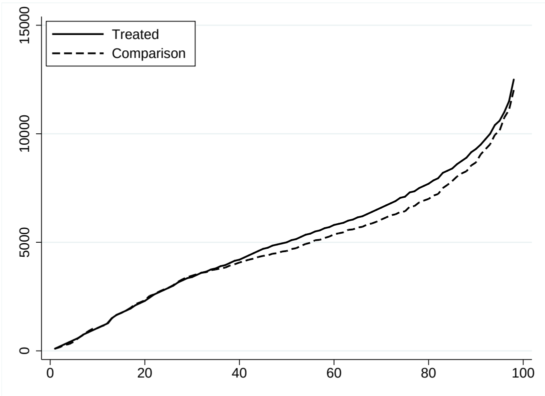
comparison individuals.[11] Panel B draws the difference between the distributions, showing that the expected heterogeneity is indeed present: Treated individuals have higher earnings at high ranks compared to untreated individuals, while the difference is small positive or negative at the lower end of the distribution.

Because Jobs First came at an extra cost compared to AFDC, this should be accounted for in our evaluation of the program. An advantage of our framework, is that it can easily accommodate alternative trade-offs between costs and benefits. To account for the surplus costs of $155 per capita for Jobs First, we introduce two hypothetical tax schedules to evaluate the cost efficiency.[12] In panel C of Figure 5.2, we subtract a lump sum tax from all participants equal to the per capita cost of the program, shifting the estimated gains down by an equal amount ($155) at every percentile. In panel D, we instead subtract a proportional income tax, defined from the relevant outcome variable. This shifts the estimated gains at every percentile down by an amount proportional to the income at that percentile ($\$155 \cdot y_1^q/\mu_1$), ensuring that the total amount collected equals the total cost. Notice that the estimated mean gains from the program of just under $300 are higher than the estimated costs per capita.

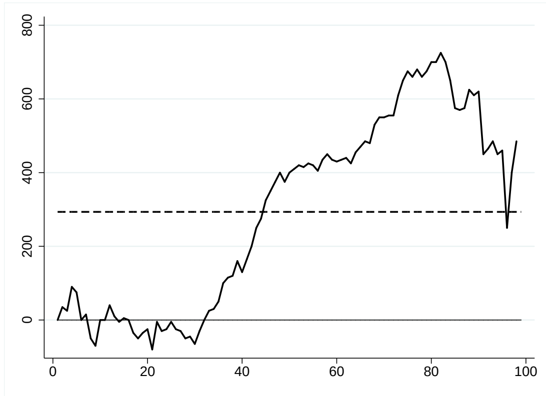
In panel A of Table 5.1, we report the estimated degree and direction of upwards and downwards inverse stochastic dominance, under the alternative hypothetical tax regimes. Remember

[11] Following Bitler, Gelbach, and Hoynes (2006), we account for observables using propensity score weighting. Note that Bitler et al. (in the published version) emphasize effects on earnings, transfers, and total income individually, and considers quarters 1–7 and 8–16 separately. Because social welfare should arguably depend on total income, and because we do not want to take a stand on consumption smoothing, we have chosen to present estimates for the full period jointly. These results also best demonstrates the application of our framework. We have also considered effects on income per person–quarter (not averaged over the period), which generates a similar picture.

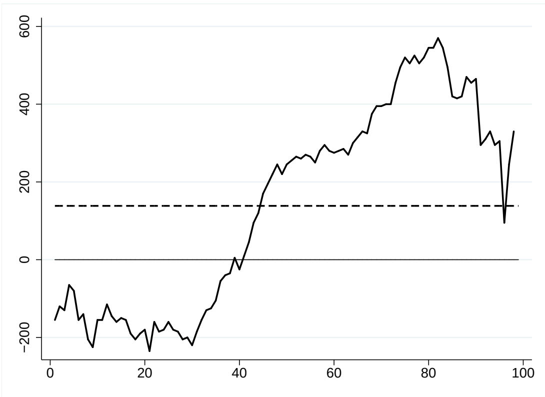
[12] As usual in such analysis, we abstract from labor supply responses.



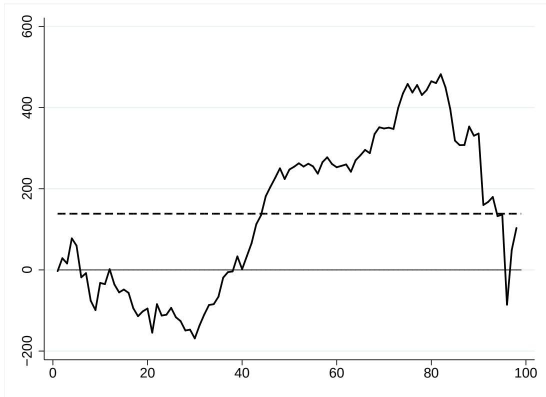
(a) Total income per person



(b) Quantile treatment effects



(c) QTE, flat tax



(d) QTE, proportional tax

Figure 5.2: Total income of treated and comparison group averaged over quarters 1–16 and estimated quantile and mean treatment effects with and without tax.

	Upward dominance			Downward dominance		
	No tax	*Lump-sum*	*Proportional*[a]	*No tax*	*Lump-sum*	*Proportional*
Inverse stochastic dominance						
Degree	4	4	231	3	3	3
Curve	1	0	0	1	1	1
ΔW_P	8.77 %	-6.16 %	–	10.94 %	0.60 %	3.85 %
$W(F_0)$	$341	$341	–	$742	$742	$742
Social welfare functions: Weights at quantiles relative to median						
p(.05)	3.61	3.61	7E+63	1.90	1.90	1.90
p(.30)	1.96	1.96	3E+33	1.40	1.40	1.40
p(.70)	0.36	0.36	2E-51	0.60	0.60	0.60
p(.95)	0.01	0.01	1E-229	0.10	0.10	0.10

Table 5.1: Estimated inverse stochastic dominance degrees and associated social welfare weights, from effects on total income averaged over quarters 1–16.
Note: Estimated effects are approximated from Bitler, Gelbach, and Hoynes (2006).

a. Estimated social welfare omitted due to rounding error that, by definition, upward dominance ranks with the poorest individual (with a non-zero difference), while downward dominance ranks with the mean. Panel B of the table shows the implied least-restrictive social welfare-weights at selected percentiles relative to the weight at the median from the estimated dominance degree.

Starting with the regime with no tax, i.e. before accounting for costs, the distribution in the comparison group 4th-degree upwards inverse stochastic dominates the distribution in the comparison group. Before accounting for costs, the estimated dominance degree implies that a social planner that is substantially averse to inequality on the downside can unambiguously rule in favor of the program. Specifically, compared to the median, a social planner that assigns at least 3.6 times the weight to individuals at the 5th percentile and about 2 times the weight to individuals at the 30th percentile can unambiguously say that Jobs First is welfare improving. In this case, W_{G_4} suggests that the program increased social welfare by 8.8 %. Notice that the relatively strong requirement on downside inequality aversion comes from the need to overtake the negative effects before the final crossing (around the 35th percentile) with positive effects at lower ranks, in order to satisfy upwards dominance.

Turning to downward dominance, the positive mean impacts imply that the program may be evaluated positively if downside inequality aversion is sufficiently weak. Specifically, the income distribution in the treatment group 3rd-degree downard inverse stochastic dominates the income distribution in the comparison group, for an implied social welfare gain of almost 11 %. Therefore, any social planner which is less averse to downside inequality than implied by the weights in panel B will rank in favor of the program. For instance, social planners that, compared to the median, assign less than 90 % and less than 40 % more weight to the 5th and 30th percentiles, respectively, will rank the program as welfare improving. Before accounting for costs, the program thus looks highly beneficial for social planners that are not too averse to downside inequality.

When we account for the costs of Jobs First, the program of course compares less favorably. Because both lump-sum and proportional taxes are larger than the gains at the lower end of the distribution, the criteria of upwards dominance now ranks the treatment below comparison. Indeed, under lump-sum taxation, comparison 4th-degree upwards dominates treatment, implying that all social welfare functions with stronger *downside* inequality aversion that could unambiguously rank the treatment as welfare improving before accounting for costs, will now rank it as unambiguously detrimental to social welfare. Under proportional taxation, however, comparison only upward dominates of degree 231: Losses at the lower end are sufficiently small that essentially only social planners employing the mini-max criterion will turn down the program unambiguously.[13]

[13]The direction of upwards dominance turns on the income difference of the poorest individuals in the two distributions. Under proportional taxation, only the 1st percentile is hit by a marginal negative effect, while effects are positive (and more sizeable) at percentiles 2–5. This illustrates how sensitive the criteria of upwards dominance may be to perturbations at the lower end. As a robustness check, nulling the negative effect at the 1st percentile, the treatment distribution 19th degree upwards inverse stochastic dominates the comparison distribution. It remains,

Meanwhile, the criteria of downward dominance still suggest a positive overall impact of the program, with the income distribution among the treated also in this case 3rd-degree downward dominating the control distribution under both tax regimes. However, while lump-sum taxation implies only a small gain in social welfare of 0.6 %, proportional taxation preserves gains at a substantial 3.9 %.

Overall, the evaluation of Jobs First when taking account of its cost, thus turns on the profile of inequality aversion in your social welfare function: Taking proportional taxation as our baseline, social planners that are not too inequality averse, assigning to the poorest individual no more than twice the weight assigned to the median, will find the program unambiguously welfare improving. Social planners that are more inequality averse than this will not be able to unambiguously rank the program against the alternative.

5.2 The evolution of the earnings distribution

6 Concluding remarks

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.

AABERGE, R., AND U. COLOMBINO (2010): "Using a Microeconometric Model of Household Labour Supply to Design Optimal Income Taxes," Carlo Alberto Notebooks 157, Collegio Carlo Alberto.

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

ATKINSON, A. B., 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.

therefore, that only social planners with very strong downside inequality aversion will be able to unambiguously rank the distributions.

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., M. BREWER, P. HAAN, AND A. SHEPHARD (2009): "Optimal Income Taxation of Lone Mothers: An Empirical Comparison of the UK and Germany," *Economic Journal*, 119(535), F101–F121.

CARNEIRO, P., J. J. HECKMAN, AND E. VYTLACIL (2010): "Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin," *Econometrica*, 78(1), 377–394.

DINARDO, J., N. M. FORTIN, AND T. LEMIEUX (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–44.

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

FORTIN, N., T. LEMIEUX, AND S. FIRPO (2011): *Decomposition Methods in Economics*vol. 4 of *Handbook of Labor Economics*, chap. 1, pp. 1–102. Elsevier.

GOTTSCHALK, P., AND T. M. SMEEDING (1997): "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature*, 35(2), 633–687.

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.

HECKMAN, J. J., AND E. VYTLACIL (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73(3), 669–738.

JUHN, C., K. M. MURPHY, AND B. PIERCE (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101(3), 410–42.

KOLM, S.-C. (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.

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

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.

A Proofs

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 an continuous function on* $[0,1]$. *Then*

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

implies

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

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

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

Proof. **Theorem 2.1.**[14] Using integration by parts and inserting $\Lambda_F^2(u)$ and $\Lambda_F^3(u)$ from Equations (2.10) and (2.11), we get that

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

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

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

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

To prove the equivalence between (ii) and (iii), consider a case where we transfer a small amount γ from persons with incomes $F^{-1}(s+h_1)$ and $F^{-1}(t+h_1)$ to persons with incomes $F^{-1}(s)$ and $F^{-1}(t)$, respectively, where $t > s$. Then W_P defined by (2) 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 $P'''(s) > 0$. □

[14] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Moreover, let

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

and

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

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

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

Inserting for (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 for $t = s + h$, 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 for (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 for (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. **Theorem 2.4.** The proof follows exactly the reasoning used in the proof of Theorem 2.3, using the following expression,

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

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

Proof. **Theorem 4.2.** Let

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

and note that

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

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

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

where

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

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

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

where the constant M is independent of N.

Since for normally distributed random variables with mean 0,

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

we have

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

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

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

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

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

converges in distribution to the process

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

□

Proof. **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 $\sum_{j=1}^{\infty}\frac{\sqrt{2}}{j\pi}\left[\int_u^1(t-u)^{i-2}\frac{\sin(j\pi t)}{f\left(F^{-1}(t)\right)}dt\right]Z_j$. □