DISCUSSION PAPER SERIES

IZA DP No. 8503

Equalization of Opportunity: Definitions, Implementable Conditions and Application to Early-Childhood Policy Evaluation

Francesco Andreoli
Tarjei Havnes
Arnaud Lefranc

September 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Equalization of Opportunity: Definitions, Implementable Conditions and Application to Early-Childhood Policy Evaluation

Francesco Andreoli
CEPS/INSTEAD and University of Verona

Tarjei Havnes
University of Oslo and IZA

Arnaud Lefranc
THEMA, University of Cergy-Pontoise and IZA

Discussion Paper No. 8503
September 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

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

IZA Discussion Paper No. 8503
September 2014

ABSTRACT

Equalization of Opportunity: Definitions, Implementable Conditions and Application to Early-Childhood Policy Evaluation*

This paper develops a criterion to assess equalization of opportunity that is consistent with theoretical views of equality of opportunity. We characterize inequality of opportunity as a situation where some groups in society enjoy an illegitimate advantage. In this context, equalization of opportunity requires that the extent of the illegitimate advantage enjoyed by the privileged groups falls. Robustness requires that this judgement be supported by the broadest class of individual preferences. We formalize this criterion in a decision theoretic framework, and derive an empirical condition for equalization of opportunity based on observed opportunity distributions. The criterion is used to assess the effectiveness of child care at equalizing opportunity among children, using quantile treatment effects estimates of a major child care reform in Norway.

JEL Classification: D63, J62, C14, I24

Keywords: equality of opportunity, public policy, inverse stochastic dominance, economic distance, income distribution, child care, pre-school

Corresponding author:

Francesco Andreoli
CEPS/INSTEAD
3 avenue de la Fonte
L-4364 Esch-sur-Alzette
Luxembourg
E-mail: francescondrl@gmail.com

* We are indebted to Claudio Zoli for valuable comments and discussions. We are also grateful to Rolf Aaberge, Marc Fleurbaey, Magne Mogstad, Matias Nuñez, Erwin Ooghe, Eugenio Peluso, Giuseppe Pignataro, Alain Trannoy, Dirk Van de gaer, Felix Weinhardt and participants at the LAGV#11 conference (Marseille 2012), 11th SSCW (New Delhi 2012), EEA-ESEM (Malaga 2012), GRASS V (Roma 2012), ECINEQ (Bari 2013), IARIW 2014, SITE Inequality Conference, and seminars at THEMA, Verona, CEPS, KU Leuven, UA Barcelona, FU Berlin for discussion. We also thank Statistics Norway for access to data and for financial support. This research received financial support from the French National Research Agency, under the grant TRANSINEQ (ANR-08-JCJC-0098-01) and FNR Luxembourg (AFR grant 5932132). Havnes also received support from The Norwegian Research Council (NRC, grant no. 212305) and is hosted by the ESOP centre funded under NRC grant no. 179552.

IZA Discussion Paper No. 8503
September 2014

NON-TECHNICAL SUMMARY

Public policies often explicitly seek to reduce inequality of opportunity. However, economic analysis lacks analytical tools to assess equalization of opportunities. This paper develops a criterion to assess equalization of opportunity that is consistent with theoretical views of equality of opportunity. In essence, to claim that a policy intervention equalizes opportunity, this criterion requires that the extent of unfair advantage enjoyed by privileged groups in society falls compared to under-privileged ones. A necessary condition for this to happen is that the economic distance between the opportunity sets of the different groups in society falls. We refer to this condition as opportunity set gap dominance. If this condition is satisfied, individuals in society will unanimously agree that inequality of opportunity has fallen.

We apply this framework to assess the effectiveness of child care at equalizing opportunity among children, in the case of a major child care expansion that took place in Norway in the 1970s. Our results indicate that the effect of expanding child care access where very heterogenous across and within family backgrounds. Children from lower family backgrounds generally benefited from the reform. This differs from more advantaged family backgrounds, where the child care expansion was not unanimously beneficial to children. In these groups, some children exhibited small gains while others experienced sizable losses. Overall, our criterion indicates that inequality of opportunity was reduced as a result of the Norwegian child care expansion.

1 Introduction

An important goal for public policy is to promote equality of opportunity, to let individual success be determined by merit rather than by social background. For instance, US President Barack Obama emphasized in his 2013 State of the Union-address the need to "make sure none of our children start the race of life already behind", in the context of policies to promote early childhood education.[1] Assessing whether public intervention succeeds at leveling the playing field among citizens thus represents a key issue for policy evaluation. But what criterion should be used to conduct such an evaluation? Unfortunately, while an abundant literature has been devoted to define *equality* of opportunity, it offers little guidance for assessing how far a given distribution is from the equality of opportunity goal. The contribution of this paper is to define a theoretical criterion of *equalization* of opportunity, understood as a reduction in the extent of inequality of opportunity and to apply this criterion to policy evaluation.

Theories of equality of opportunity (EOP) draw a distinction between fair inequality, arising from differences in individual *effort*, and unfair inequality arising from differences in individual *circumstances*, i.e. the determinants of success for which society deems the individual not to be responsible.[2] Define a *type* as a given set of circumstances, and an *opportunity set* as the set of feasible outcomes for each type. The EOP-principle requires that no type is advantaged compared to other types in the sense of having access to a more favorable opportunity set. This principle allows to assess whether a given distribution satisfies equality of opportunity. However, it does not allow to compare two societies where equality of opportunity is not satisfied. This is obviously an important limitation in many contexts, including policy evaluation and comparisons of inequality across time and space.

To alleviate this shortcoming, some authors have relied on indices of inequality of opportunity.[3] This approach typically seeks to produce a scalar measure of inequality between types.[4] While consistent with the EOP principles, this approach raises concerns

[1] State of the Union Address, February 12, 2013.

[2] For a comprehensive discussion, see Dworkin (1981), Roemer (1998) and Fleurbaey (2008).

[3] See the discussion in Ramos and Van de gaer (2012), and examples in Bourguignon, Ferreira and Menendez (2007), Lefranc, Pistolesi and Trannoy (2008), Checchi and Peragine (2010), Aaberge, Mogstad and Peragine (2011), Almås, Cappelen, Lind, Sørensen and Tungodden (2011), Björklund, Jäntti and Roemer (2011) and Ferreira and Gignoux (2011).

[4] Checchi and Peragine (2010) and Ramos and Van de gaer (2012) discuss a dual approach where inequality of opportunity is measured by the difference between total inequality and fair inequality. The two limitations discussed here also apply to the dual approach.

of robustness as it relies on two restrictive assumptions. First, it requires summarizing the advantage enjoyed by a type in a scalar measure. Usually the mean income conditional on circumstances is used to summarize the opportunity sets faced by each type.[5] But these scalar measures may mask important features of the distribution of opportunity. Second, inequality index approaches rely on a specific welfare function used to aggregate differences in advantage between types. They are therefore fragile to the precise specification of the welfare function, including both parameter choices and functional form.

In this paper, we develop a robust criterion for *equalization of opportunity* (EZOP) that allows statements such as: "Inequality of opportunity is higher in social state 0 than in social state 1," where different states might correspond to different countries, time periods or policy regimes. Our criterion has an appealing interpretation, and is readily implementable in policy evaluation. The criterion requires that individuals, regardless of their preferences, agree that the advantage enjoyed by the "privileged" types is lower in policy state 0 than in policy state 1.

Contrary to the index approach, our criterion does not rely on a priori value functions to assess the advantage enjoyed by each type. Instead, we use the preferences over opportunity sets of individuals in society and allow for heterogeneity in these preferences. Robustness requires a consensus across individuals in their comparison of social states. A key question in this respect is whether such a consensus can be reached. When consensus cannot be reached, an important issue is to characterize the subset of preferences over which individuals unanimously agree that equalization of opportunity is achieved. The equalization of opportunity criterion is also demanding in requiring that the unfair advantage of "privileged" types falls for any pairwise comparison of types. We discuss how this criterion can be relaxed by allowing the advantage of each type to be aggregated within society, for given individual preferences.

Our criterion also raises an important issue of identification. In practice, we only observe (at best) the opportunity sets of each type but we do not observe individual preferences. Hence it is not feasible to verify for each particular preference whether the advantage of privileged types is lower in one particular state. Instead, we would like to define a tractable condition, involving only the distribution of opportunity sets, that would imply that our equalization criterion is satisfied. We show that such a condition

[5]This amounts to assume that individuals are risk neutral, with respect to within type uncertainty. Lefranc et al. (2008) assume risk aversion but rely on specific preferences.

can be formulated provided that individuals agree on the ranking of types in each social state. On the contrary, when individuals disagree on the ranking of types, they cannot unanimously agree on equalization of opportunity. However, in this case, it is possible to identify subclasses of preferences within which individuals agree on the ranking of types in each state and to single out a necessary and sufficient condition for equalization within this subclass of preferences. This can only be performed within a specified set of preferences. In this paper, we focus mainly on the rank-dependent representation of the preferences (Yaari 1987), although the analysis can be adapted to other classes.[6] Another important issue in empirical applications, is that some relevant determinants of individual outcomes might not be fully observed.[7] For instance, circumstances or effort might be only partially observed. We discuss the consequences of imperfectly observing the relevant determinants of outcome for the implementation of our equalization criterion.

Finally, we show the usefulness of our framework by applying it to the evaluation of child care policy in Norway. We follow Havnes and Mogstad (2011, 2014) in considering how the introduction of universally available child care in Norway affected children's adult earnings. We use a similar identification strategy to estimate quantile treatment effects, exploiting the spatial and temporal variation of the expansion in a difference-in-differences setup. We extend on their framework to allow impacts across the distribution to vary flexibly with family background. Overall, results confirm the substantial heterogeneity documented in Havnes and Mogstad (2014). Our results complement theirs, by documenting that impacts were in fact quite homogenous across family backgrounds in the lower end of the distribution, with positive effects, but highly dispersed in the upper end, where children from disadvantaged groups gained while children from advantaged groups lost from the expansion. When applied to our EZOP framework, the results suggest that the child care expansion significantly equalized opportunities between children from most family backgrounds. At the same time, the result is driven importantly by losses among upper-class children, and equalization is mostly not achieved when we consider the most disadvantaged group who experience relatively small gains.

[6]The relevance of the rank-dependent representation of preferences in income inequality analysis has been documented in several papers (Muliere and Scarsini 1989, Maccheroni, Muliere and Zoli 2005, Aaberge 2009, Aaberge and Mogstad 2011). Aaberge, Havnes and Mogstad (2014) propose a robust welfarist criterion for ranking income distributions, based on unanimous agreement between subclasses of evaluation functions admitting the rank-dependent representation. This paper adapts and extends some of these results to the equality of opportunity framework.

[7]See Dardanoni, Fields, Roemer and Sanchez Puerta (2005) for a discussion. Also see Björklund et al. (2011) for a comprehensive measure of individual circumstances

The rest of the paper is organized as follows. Section 2 introduces the principle of equalization of opportunity, and provides a tractable condition in a simplified setting with two types. Next, section 3 considers the general case with multiple types, effort and circumstances. Finally, section 4 applies our framework in the context of a major expansion in child care in Norway. As an addendum, we also develop a statistical framework for testing our equalization condition, which is included in appendix B.

2 Equalization of opportunity: a simplified setting

In this section, we define equality of opportunity and provide a formal statement of our equalization criterion in a simplified setting. Next, we discuss identification conditions that guarantee that this criterion is satisfied.

2.1 Definition of equality of opportunity

Our analysis builds upon the framework of Roemer (1998) and Lefranc, Pistolesi and Trannoy (2009). Let y denote an individual outcome, and let the determinants of the outcome be partitioned into four groups: *Circumstances* capture determinants that are not considered legitimate sources of inequality, and are denoted by c. *Effort* captures determinants that are considered legitimate sources of inequality, and is denoted by e. *Luck* captures factors that are considered legitimate sources of inequality as long as they affect individual outcomes in a neutral way given circumstances and effort, and is denoted by l. Finally, outcomes are contingent on a binary social state, denoted π. All individuals in a society share the social state, but may be affected differently. For instance, $\pi = 0$ may denote society without a specific policy intervention, while $\pi = 1$ denotes society with the policy, or π may indicate different periods or countries that one would like to compare.

Let a *type* define a given set of circumstances. Given their type, level of effort and the social state, the outcome prospects offered to individuals can be summarized by the cumulative distribution function $F_\pi(y|c,e)$. We define $F_\pi^{-1}(p|c,e)$ as the conditional quantile function associated with $F_\pi(\cdot|c,e)$, for all population shares p in [0,1].[8]

EOP theories emphasize that inequality due to differences in circumstances are morally or politically objectionable, while inequality originating from differences in effort are legit-

[8] If the cumulative distribution function is only left continuous, we define F_π^{-1} by the left continuous inverse distribution of F_π: $F_\pi^{-1}(p|c,e) = \inf\{y \in \mathbb{R}_+ : \ F_\pi(y|c,e) \geq p\}$, with $p \in [0,1]$.

imate. Based on these two principles, *equality of opportunity* requires that the opportunity sets of individuals with similar effort be *identical* regardless of circumstances. Hence, for a given social state π, EOP requires that, for any effort e, and for any pair of circumstances (c, c'), we have:

$$F_\pi(.|c, e) \;=\; F_\pi(.|c', e) \tag{1}$$

2.2 A criterion for equalization of opportunity

Ranking social states The previous definition can be used to rank social states. However, it only distinguishes states where EOP is satisfied and states where EOP is not satisfied. Since the strict equality in equation (1) is rarely satisfied in empirical applications, this leads to a very partial ranking.[9]

Our objective is to provide a criterion that allows to compare and rank social states in situations where EOP is not satisfied. If EOP is not satisfied, then individuals are not indifferent between the opportunity sets offered to different types. Behind a thin veil of ignorance, where individuals know their effort and have preferences over opportunity sets, everyone should indeed be able to rank circumstances according to the economic advantage or disadvantage they confer. Our criterion for ranking social states is based on the evaluation of the extent of the economic advantage enjoyed by the advantaged types in society. It posits that social state 1 is better than social state 0, in terms of equality of opportunity, if the unfair advantage attached to favorable circumstances is lower in state 1 than in state 0. When preferences are heterogenous, individuals might not agree on the ranking of social states. To ensure robustness, our equalization of opportunity criterion (EZOP) requires unanimity *for all possible preferences in society*, in assessing that the unfair advantage attached to more favorable circumstances decreases.

For expositional purposes, we start by formalizing the equalization criterion in a simplified setting with only two types, c and c', who exert a common effort level e. To simplify notation, we let $F_\pi(.)$ (resp. $F'_\pi(.)$) denote the c.d.f. of y for type c (resp. c') at effort e in policy state π, *i.e.* $F_\pi(.|c, e)$ (resp. $F_\pi(.|c', e)$). Section 3 provides a generalization with many types and effort levels.

[9] For a survey of the empirical evidence, see Roemer and Trannoy (2014).

The EZOP criterion We assume that each individual is endowed with cardinal preferences over risky outcomes and we let $W(F)$ denote the utility of a lottery with cumulative distribution F. For an individual with preferences W, the economic advantage or disadvantage of type c relative to type c', in social state π is denoted $\Delta_W(F_\pi, F'_\pi)$ and is given by $\Delta_W(F_\pi, F'_\pi) = W(F_\pi) - W(F'_\pi)$. This quantity is positive if the individual with preferences W prefers F_π to F'_π, while it is equal to zero if EOP holds between types c and c'. We refer to the absolute value of the welfare gap as the *economic distance between types* according to preferences W.[10]

The equalization of opportunity criterion rests on the difference in economic advantage across social states. Let $\mathcal{P}$ denote the class of individual preferences. The following definition summarizes our notion of equalization of opportunity:

Definition 1 (EZOP: equalization of opportunity between two types) *Moving from policy state $\pi = 0$ to $\pi = 1$ equalizes opportunity between circumstances c and c' at effort e on the set of preferences $\mathcal{P}$ if and only if for all preferences $W \in \mathcal{P}$, we have:* $|\Delta_W(F_0, F'_0)| \geq |\Delta_W(F_1, F'_1)|$.

This notion has three key properties. First, in line with the theory of EOP, to rank state 1 above state 0, we require that the unfair economic advantage enjoyed by the privileged type be smaller in state 1 than in state 0. Second, our criterion satisfies an anonymity condition with respect of the identity of the advantaged type: Only the absolute value of the economic advantage, but not its sign, should matter for assessing equalization of opportunity. Third, we require that the ranking is robust to a broad class of individual preferences.

2.3 Identification under the rank-dependent utility model

The identification problem The EZOP criterion is contingent on the choice of the class of preferences $\mathcal{P}$. If the set of individual preferences W in society were known, the economic distance between types, Δ_W, could be directly computed for both social states and it would be easy to check whether the equalization condition holds. In practice, preferences are not known and the condition in definition 1 cannot be directly assessed for all relevant preferences.

[10] For a discussion of the welfare gap and related measures of economic distance between distributions, see in particular Shorrocks (1982), Ebert (1984) and Chakravarty and Dutta (1987).

To make the equalization condition previously defined relevant, we need to reformulate it in terms of a restriction that involves only the outcome distributions of the different types in the alternative states. This cannot be achieved in the most general case where no restriction is imposed on the class $\mathcal{P}$ of individual preferences. Two possible alternative representations of preferences under risk have been widely studied and adopted in decision theory: the *expected utility* model and the *rank-dependent* model of Yaari (1987). In the rest of the paper, we focus on the class of preferences in the rank-dependent family, which we denote by $\mathcal{R}$. In the rest of this section we concentrate on the following question: What minimal conditions need to be imposed on the set of distributions F_0, F_0', F_1, F_1' to ensure that equalization is satisfied for all preferences in $\mathcal{R}$?

The rank-dependent model assumes that the welfare derived from a risky distribution F can be written as a weighted average of all possible realizations, where the weights are a function of the rank of the realization in the distribution of outcome. Formally, let $w(p) \geq 0$ denote the weight assigned to the outcome at percentile p, the welfare derived from F can be written as[11]:

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

Under the rank-dependent representation, the economic distance between types is given by:

$$|\Delta_W(F, F')| = |\int_0^1 w(p)\Gamma(F, F', p)dp|. \tag{2}$$

where $\Gamma(F, F', p) = F^{-1}(p) - F'^{-1}(p)$ is the *cumulative distribution gap* between F and F'. In the rest of the paper, we refer to the graph of $\Gamma(F, F', p)$ as the *gap curve* and to the graph of $|\Gamma(F, F', p)|$ as the *absolute gap curve.*

Necessary condition for EZOP When assessing EZOP, only the distribution curves under the two policy states are observed, but not individual preferences. Our objective is to provide a condition on these observables warranting that EZOP holds for all preferences in the class of rank-dependent utility functions. From equation (2), a necessary condition for EZOP is that the cumulative distribution gap under $\pi = 1$ should be smaller, in absolute value, compared to the gap under $\pi = 0$, at all percentiles.

[11] Formally, one requires that $w(p) \geq 0 \; \forall p \in [0, 1]$ and $\widetilde{w}(p) = \int_0^p w(t)dt \in [0, 1]$ is such that $\widetilde{w}(1) = 1$. For a discussion, see Zoli (2002).

Proposition 1 *EZOP is satisfied on the set of preferences* $\mathcal{R} \Rightarrow \forall p \in [0,1]$, $|\Gamma(F_1, F_1', p)| \leq |\Gamma(F_0, F_0', p)|$.

Proof. See appendix A.2. ■

This proposition shows that a necessary condition for EZOP is that the absolute gap curve under $\pi = 0$ is always above the absolute gap curve under $\pi = 1$. We will refer to this as absolute gap curve dominance of $\pi = 0$ over $\pi = 1$.

Note that absolute gap curve dominance is *not* a *sufficient* condition for EZOP.[12] Whether a reduction in the gap between type c and c' amounts to a reduction in advantage, will depend on which of the two groups is considered to be advantaged. Because the assessment of which type is advantaged may differ over the set of possible preferences, the requirement for EZOP over all possible preferences must be stronger than what is imposed by absolute gap curve dominance. For instance, assume that the distribution of type c dominates the distribution of type c' over some interval. This does not imply, in the general case, that type c dominates c' over the entire support of the distribution. Henceforth, some preferences might rank c' better than c. Now assume that gap curve dominance is satisfied over this interval and that gap curves are similar in both social states otherwise. In this case, preferences that ranked c' better than c will conclude that the cardinal advantage of c' has increased. This contradicts EZOP.

Necessary and sufficient condition under stochastic dominance A corollary of the previous discussion is that if individuals agree on the ranking of types, then they should also agree in their assessment of gap curve dominance. We now examine this specific case.

As discussed in Muliere and Scarsini (1989), among others, unanimity in ranking distributions F_π better than F_π' will be achieved for all preferences in $\mathcal{R}$ if and only if distribution F_π stochastically dominates distribution F_π' (which we denote $F_\pi \succ_{ISD1} F_\pi'$ for consistency with the below), *i.e.* whenever the graph of F_π lies above the graph of F_π'.[13]

In this section, we assume that this condition is satisfied.[14] If so, all preferences

[12] Appendix A.2 provides a detailed counter-example.

[13] Note that stochastic dominance and inverse stochastic dominance are equivalent on the first and second order. The difference is that the dominance condition in the latter case is expressed in the space of realizations (through the quantile function) while in the former case it is expressed in the space of probabilities (through the cdf).

[14] Since c and c' play a symmetric role in the definition of EZOP, which type dominates the other is irrelevant. Hence we make the neutral assumption that the distribution of type c dominates the distribution of type c', under both policy regimes.

unanimously rank type c better than type c'. A fall in the cumulative distribution gap then has unambiguous consequences for the change in the economic distance between types. In fact, since the sign of the cumulative distribution gap is constant across all percentiles, the economic distance can be expressed as an increasing function of the absolute income gap: $|\Delta_W(F, F')| = \int_0^1 w(p)|\Gamma(F, F', p)|dp$. This leads to the following proposition:

Proposition 2 *If* $\forall \pi \ F_\pi \succ_{ISD1} F'_\pi$ *then: EZOP over the set of preferences* $\mathcal{R} \ \Leftrightarrow \ \forall p \in [0,1], \ \Gamma(F_0, F'_0, p) \geq \Gamma(F_1, F'_1, p)$.

Proof. See appendix A.3. ■

This proposition establishes that when agents agree on the ranking of types, gap curve dominance provides a necessary and sufficient condition for equalization of opportunity. This contrasts with the situation where agents do not agree on the ranking of types, in which case gap curve dominance provides only a necessary condition for EZOP. In order to evaluate EZOP in such situations, we next consider refinements on the admissible set of preferences.

Restricted consensus on EZOP We now focus on cases where types cannot be ranked unambiguously, i.e. according to first-order stochastic dominance. As we have seen, the cumulative distribution gap is then no longer sufficient to infer EZOP. Our objective is to identify the minimal refinement on the set of admissible preferences that allow unambiguous assessments of equalization of opportunity. In line with Aaberge et al. (2014), we show that it is always possible to find a subset of $\mathcal{R}$ over which individuals agree on the ranking of types. Furthermore, on this subset, one can establish a necessary and sufficient condition for equalization of opportunity.

Let us first consider the special case where F_π *second-order* stochastic dominates F'_π for all $\pi \in \{0,1\}$ (which we denote $F_\pi \succ_{ISD2} F'_\pi$, again to stay consistent with the notation below), *i.e.* whenever the graph of the integral of F_π^{-1} with respect to p (the Generalized Lorenz curve) lies above the graph of the corresponding integral of $F_\pi'^{-1}$.[15] Define $\mathcal{R}^2 \subset \mathcal{R}$ as the set of risk-averse rank-dependent preferences.[16] As is well known, all risk averse preferences rank distribution functions consistently with second-order dominance. It follows

[15] Note that stochastic dominance and inverse stochastic dominance are equivalent on the second order, and both are equivalent to generalized Lorenz dominance.

[16] This set contains all evaluation functions with decreasing weights as outcomes increase, i.e. that have $w'(p) < 0$.

that all preferences in $\mathcal{R}^2$ will rank type c as better than c' in both states. Furthermore, the advantage of c over c' can be expressed as an increasing function of *the integral* of the cumulative distribution gap. Analogous to the above, a necessary and sufficient condition for EZOP over the set of preferences $\mathcal{R}^2$ is then that the integrated cumulative distribution gap falls at all percentiles. This is established in the following proposition:

Proposition 3 *If* $\forall\pi\ F_\pi \succ_{ISD2} F'_\pi$ *then: EZOP over the set of preferences* $\mathcal{R}^2 \ \Leftrightarrow\ \forall p \in [0,1], \int_0^p \Gamma(F_0, F'_0, t)dt \ \geq \int_0^p \Gamma(F_1, F'_1, t)dt$

Proof. See appendix A.4. ■

Finally, consider the general case where distributions cannot be ranked by second-order stochastic dominance. In this case, consensus over the ranking of types cannot be reached in the class $\mathcal{R}^2$. However, it is possible to refine the set of preferences to where they agree on the ranking of types. Following Aaberge (2009), consider the subset of preferences $\mathcal{R}^k$ defined by:

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

where $\widetilde{w}(p) = \int_0^p w(t)dt$ is the cumulative weighting scheme. The sequence of subsets of the type $\mathcal{R}^k$ defines a nested partition of $\mathcal{R}$ where $\mathcal{R}^k \subset \mathcal{R}^{k-1} \subset ... \subset \mathcal{R}$.[17]

Various papers have examined the relationship between inverse stochastic dominance and the ordering of distributions according to preferences in $\mathcal{R}^k$ (Muliere and Scarsini 1989, Zoli 2002, Aaberge 2009). Aaberge et al. (2014) provide a general treatment and show that for any order k all preferences in $\mathcal{R}^k$ will prefer F_π over F'_π if and only if F_π inverse stochastic dominates F'_π at order k. Furthermore, as we show in appendix A.1, any pair of distributions can always be ranked by inverse stochastic dominance, for a sufficiently high finite order.[18] Define κ as the *minimal order* at which F_π and F'_π can be ranked using inverse stochastic dominance in both states, and denote k^{th} order inverse stochastic dominance by $\succ_{ISDk}$. Without loss of generality, assume that $F_\pi \succ_{ISD\kappa} F'_\pi$ for all $\pi \in \{0,1\}$, such that preferences in $\mathcal{R}^\kappa$ agree on the ranking of types in both states.

[17] Note that k is a measure of the effect of a precise sequence of restrictions on all possible cumulative weighting schemes $\widetilde{w}(p)$ defined on $\mathcal{R}$. Hence, k indicates the risk attitude of preferences contained in the class $\mathcal{R}^k$.

[18] This result holds for distributions with bounded support. It is connected to the result in Aaberge et al. (2014) that, as k goes to infinity, order k inverse stochastic dominance comparisons amount to a comparison of the lowest income in each distribution with bounded support.

To proceed, it is helpful to introduce the following notation:

$$\Lambda_\pi^2(p) = \int_0^p F_\pi^{-1}(u)du \quad \text{and} \quad \Lambda_\pi^k(p) = \int_0^p \Lambda_\pi^{k-1}(u)du, \text{ for } k = 3, 4, \ldots$$

For notational simplicity, we let $\Lambda_\pi'^k$ denote Λ_π^k evaluated over the distribution F_π' rather than F_π. In line with the notation above, also define $\Gamma^k(F_\pi, F_\pi'^k, p) = \Lambda_\pi^k(p) - \Lambda_\pi'^k(p)$ as the cumulative distribution gap integrated at order $k-1$.

If for all $\pi \in \{0,1\}$, $F_\pi \succ_{ISD\kappa} F_\pi'$, then for all preferences $W \in \mathcal{R}^\kappa$, the advantage of type c over type c' under policy π is an increasing function of $\Gamma^\kappa(F_\pi, F_\pi', p)$. As a consequence, EZOP will be satisfied on the set of preferences $\mathcal{R}^\kappa$ if and only if $\Gamma^\kappa(F_\pi, F_\pi', p)$ is smaller under $\pi = 1$ than under $\pi = 0$. This is established in the following proposition:

Proposition 4 *If* $\forall \pi \; F_\pi \succ_{ISD\kappa} F_\pi'$ *then: EZOP over the set of preferences* $\mathcal{R}^\kappa \;\Leftrightarrow\; \forall p \in [0,1],\; \Gamma^\kappa(F_0, F_0', p) \geq \Gamma^\kappa(F_1, F_1', p)$.

Proof. See appendix A.5. ■

Proposition 4 establishes a necessary and sufficient condition for EZOP under a less stringent dominance condition than in propositions 2 and 3. At the same time, the set of preferences over which it allows to identify EZOP is more restrictive. Finally, since there always exists an integer κ that allows ranking of types, proposition 4 establishes a necessary condition for EZOP over the entire class $\mathcal{R}$.

2.4 Discussion

Several features of our equalization criterion are worth discussing further. First, the criterion presented in definition 1 does not resort to an external social welfare function in order to evaluate the opportunity sets offered to the different types in society. On the contrary, the criterion relies on the individuals' *own* preferences in order to assess whether equalization of opportunity is achieved. The degree of heterogeneity of preferences distributions across the population is clearly unobservable. The focus is therefore on the class of *potential* preferences these individuals may have. Second, the criterion itself is general, in the sense that it does not place any restriction on the preferences of individuals. Third, the criterion does not in itself require that individuals agree on the ranking of types, only that they agree on the reduction in the absolute gap between the different types. In other words, our criterion requires a consensus on the reduction of the advantage but not on the

identity of the advantaged type. Finally, the criterion does not require one to summarize the opportunity sets of the different types by a scalar measure, such as the mean income, as is often done in the literature on the measurement of inequality of opportunity.

Of course, while the generality of the criterion leads to a robust assessment of equalization of opportunity, this robustness comes at the cost of tractability. As we noted, it is not possible, in practice, to verify whether the condition of equalization is satisfied without considering a restricted set of preferences. In the rest of our analysis, we considered the family of rank-dependent utility functions and showed that it is possible to derive equalization conditions that only depend on the distribution functions of the lotteries offered to the various types in society. However, our framework is not confined to the rank-dependent family and could be extended to other families of preferences. For instance, in the same spirit, equalization conditions could be derived for preferences within the Von Neumann expected utility framework.

The results obtained under the rank-dependent assumption also call for further comments. They lead to distinguish between two cases: the case where individuals agree on the ranking of types under each social state, and the case where they do not agree. When individuals agree on the ranking of types, assessing equalization of opportunity is straightforward, as proposition 2 provides a necessary and sufficient condition.

The case where individuals do not agree on the ranking of types does not allow such a clear cut judgment on equalization. Proposition 1 provides a necessary condition of equalization. Violation of this condition rules out equalization of opportunity. If not, proposition 4 allows to endogenously identify a restricted set of preferences over which unanimity might be reached regarding equalization of opportunity. Of course, this only provides a partial judgment over equalization of opportunity. In fact, the higher the order of κ required to successfully rank opportunity sets, the less general the judgement will be.

However, the extent of the restrictions on preferences required to achieve a consensus on the ranking of types may in itself be informative. When few restrictions are required to achieve a consistent ranking, then most individuals should agree on which type is advantaged. On the contrary, when strong restrictions are required, there may be widespread disagreement on which type is advantaged. In this case, one might argue that a weak form of equality of opportunity already prevails. Lefranc et al. (2009) introduce the notion of *weak equality of opportunity* to single out situations where the opportunity sets differ

across types but cannot be ranked according to second-order stochastic dominance. In this case, there is not a consensus on the identity of the advantaged type among agents with risk-averse preferences. By capturing the degree of consensus on the advantaged type among potential preferences, κ helps generalize the notion of weak equality of opportunity introduced in Lefranc et al. (2009).

To summarize, when there may be widespread disagreement on which type is advantaged (high κ), our criterion provides a very partial condition for consensus on equalization of opportunity, although this admittedly corresponds to a case of weak inequality of opportunity. On the contrary, when there is large agreement on which type is advantaged (low κ), our equalization condition becomes least partial and turns into a necessary and sufficient condition for EZOP in the case where there is full consensus on the identity of the advantaged type ($\kappa = 1$).

3 Equalization of opportunity: generalization

In the general case, opportunity equalization has to be assessed with more than two circumstances across many effort levels. When effort is observable, one possibility is to extend the EZOP comparisons on all pairs of circumstances at every effort level, or to study meaningful aggregations of these judgements. Identification criteria when effort is not observable are also discussed, in order to provide relevant notions of equalization that can still be used in applied analysis, under observability constraints.

3.1 Extending the EZOP criterion to multiple circumstances

We consider the case in which there are T types. Let $C = \{c_1, ..., c_i, ...c_T\}$ denote the set of possible circumstances. For simplicity, we assume a single effort level e. The results of this section can be easily extended to multiple effort levels by requiring that equalization holds for every effort level.

A straightforward extension of definition 1 to multiple circumstances is to require that for every possible pair of circumstances, the distance falls when moving from social state $\pi = 0$ to $\pi = 1$. This is given by the following definition:

Definition 2 (Non-anonymous EZOP between multiple types) *Moving from state $\pi = 0$ to $\pi = 1$ equalizes opportunity over the set of circumstances C at effort e on the set*

of preferences $\mathcal{P}$ if and only if for all preferences $W \in \mathcal{P}$, for all $(i,j) \in \{1,\ldots,T\}$, we have: $|\Delta_W(F_0(.|c_i,e),F_0(.|c_j,e))| \geq |\Delta_W(F_1(.|c_i,e),F_1(.|c_j,e))|$.

Again, this generalized form of EZOP cannot be verified, in practice, without resorting to a specific class of preferences. In the class $\mathcal{R}$, the results of propositions 2 and 4 generalize easily to the multivariate case. For every pair (i,j), let κ_{ij} denote the *minimal order* at which $F_\pi(.|c_i,e)$ and $F_\pi(.|c_j,e)$ can be ranked according to inverse stochastic dominance, for all π. According to proposition 4, integrated gap curve dominance for each pair of types c_i and c_j provides a necessary and sufficient condition for EZOP between the two types over the subclass $\mathcal{R}^{\kappa_{ij}}$. This condition is, however, only necessary when looking at the whole class $\mathcal{R}$.

Proposition 5 *EZOP between multiple types over the set of preferences* $\mathcal{R} \Rightarrow \forall(i,j) \in \{1,\ldots,T\}$, $\forall p \in [0,1]$, $\left|\Gamma^{\kappa_{ij}}\left(F_0(.|c_i,e),F_0(.|c_j,e),p\right)\right| \geq \left|\Gamma^{\kappa_{ij}}\left(F_1(.|c_i,e),F_1(.|c_j,e),p\right)\right|$.

The proof is based on the same arguments used in the proof of proposition 4. Differently from this proposition, while we know that $F_\pi(.|c_i,e)$ and $F_\pi(.|c_j,e)$ can be ranked at the order of dominance κ_{ij}, the direction of dominance is a priori undetermined. This explains why gap dominance should hold in absolute value.

Definition 2 makes the "identity" of each type relevant for defining equalization of opportunity, since the extent of advantaged between any pair (c_i,c_j) under $\pi = 0$ is confronted with the extent of advantage between the same two types under $\pi = 1$. One may challenge this view and claim that only the magnitude of the gaps (and not the identity of the types involved) is relevant for defining equalization of opportunity. Consider a simple example in which there are three types c_1, c_2 and c_3. Assume that there is only one effort level and luck plays no role. Under each of the three states $\pi = A, B, C$, each type is assigned with an outcome given by the following table:

	Outcomes		
	$\pi = A$	$\pi = B$	$\pi = C$
c_1	6	6	6
c_2	3	4	2.5
c_3	1	2.5	4

When moving from state A to B, the gap between each type and the other two falls (the gap between c_1 and c_2 shrinks from $6-3$ to $6-4$) and the condition in definition

2 is satisfied. On the contrary, when moving from state A to C, the gap between c_1 and c_2 increases and equalization is not satisfied, although state C is obtained from state B only by permuting the outcomes of groups c_2 and c_3. This inconsistency arises from the fact that the assessment of equalization of opportunity in definition 2 is sensitive to the identity of the groups associated to a given opportunity gap.

This counterexample echoes a well-know *anonymity principle* used in the assessment of inequality. According to this principle, the measurement of inequality of outcome should be insensitive to a permutation of the outcomes of individuals within the distribution. This principle can be incorporated in our definition of equalization of opportunity by making it insensitive to a permutation of the opportunity sets across types.

Let us introduce some additional notation. Let $r_\pi^W(c)$ be the *rank function* assigning to circumstance c its rank, $r(.) \in \{1, \ldots, T\}$ in the ranking of types, in social state π, according to preferences W. Given W, all circumstances can be ranked, but the rank of a specific circumstance c might change across social states and differs across preferences.

The anonymous principle of equalization of opportunity between multiple types requires that the opportunity gap between two types sitting at given ranks falls when moving from social state 0 to 1. This should hold for every pair of ranks and every utility function in $\mathcal{P}$.

Definition 3 (Anonymous EZOP between multiple types) *Moving from state $\pi = 0$ to $\pi = 1$ equalizes opportunity over the set of circumstances C at effort e on the set of preferences $\mathcal{P}$ if and only if for all preferences $W \in \mathcal{P}$, for all $(i, j, h, \ell) \in \{1, \ldots, T\}^4$ such that $r_0^W(c_i) = r_1^W(c_h)$ and $r_0^W(c_j) = r_1^W(c_\ell)$ we have:*
$|\Delta_W(F_0(.|c_i, e), F_0(.|c_j, e))| \geq |\Delta_W(F_1(.|c_h, e), F_1(.|c_\ell, e))|$.

Implementation of anonymous equalization requires first to define a class of preferences $\mathcal{P}$ where agreement is reached on the rank of types according to the advantage they confer, and, second, to check if the advantage between circumstances occupying a similar rank is reduced when changing state. Within the class $\mathcal{R}$, agreement on this ranking of circumstances is reached only for the evaluation functions in the intersection of all the sets $\mathcal{R}^{\kappa_{ij}}$, for all pairs (i, j). The intersection is denoted by $\mathcal{R}^{\kappa_{\max}}$, where $\kappa_{\max} = \max_{i,j \in \{1,\ldots,T\}} \{\kappa_{ij}\}$. Once this set is identified, gap curve dominance can be tested. It only provides, however, a necessary condition for agreement over $\mathcal{R}$.

Proposition 6 *EZOP between multiple types over the set of preferences* $\mathcal{R} \Rightarrow \forall p \in [0,1]$, $\forall (i,j,h,\ell) \in \{1,\ldots,T\}^4$ *such that* $r_0^W(c_i) = r_1^W(c_h)$ *and* $r_0^W(c_j) = r_1^W(c_\ell)$, $\left|\Gamma^{\kappa_{\max}}\left(F_0(.|c_i,e), F_0(.|c_j,e), p\right)\right| \geq \left|\Gamma^{\kappa_{\max}}\left(F_1(.|c_h,e), F_1(.|c_\ell,e), p\right)\right|$.

The proof is based on the same arguments used in the proof of proposition 4. Propositions 5 and 6 require that advantage gaps fall for all possible pairs of circumstances. However, some gaps might be more worth compensating than others. For instance, one might assign priority to the bottom of the distribution of types (i.e. reduce the gap between the bottom type and other types) or to the top of the distribution. This would amount, in definition 2, to restrict the scope of inter-type comparisons to pairs involving either the bottom type or the top type. This criterion remains, nevertheless, disaggregated and demanding: it requires performing a large number of comparisons of pairs of types that must all be validated by all preferences in a sufficiently heterogeneous class.

3.2 Aggregation across circumstances

It might be argued that a small increase in the opportunity gap between two types might be compensated by a fall in the opportunity gap between another pair of types. This view suggests aggregating welfare gaps $|W\left(F_\pi(.|c_i,e)\right) - W\left(F_\pi(.|c_j,e)\right)|$ across pairs of circumstances. This leads to a scalar measure of inequality of opportunity. Of course, implementing such a scalar measure requires selecting a particular preference function W. It also requires taking into account the size of the various types when aggregating welfare gaps. Define p_c as the relative frequency of type c in the population. One can define, for a function W, an *Inequality of Opportunity Indicator* (IO):

$$IO(\pi) \;=\; \sum_{i=1}^{T} \sum_{j=i+1}^{T} p_{c_i}\, p_{c_j} \left|W\left(F_\pi(.|c_i,e)\right) - W\left(F_\pi(.|c_j,e)\right)\right|.$$

IO equals the average absolute welfare gap, across all pairs of circumstances, computed for function W. This appears as a generalization of several inequality of opportunity indices suggested in the literature.[19] Lefranc et al. (2008) introduce the *Gini Opportunity* index

[19] Checchi and Peragine (2010) undertake a similar approach. They define, in a ex-post setting with degenerate luck, indicators of inequality of opportunity. Their indicators measure relative inequality among individual realizations, under the assumption that all individuals in a type exerting similar effort receive similar outcomes. Their indicator aggregates outcome differentials not only across types, but also across effort levels, something that is not necessarily imposed in the $IO(\pi)$ index. A different approach is instead undertaken in Peragine (2002, 2004), where the objects of interest are social evaluation functions and types

defined as:

$$GO(\pi) \;=\; \frac{1}{\mu}\sum_{i=1}^{T}\;\sum_{j=i+1}^{T} p_{c_i}\; p_{c_j}\big|\mu_{c_i}(1-G_{c_i}) \;-\; \mu_{c_j}(1-G_{c_j})\big|.$$

This amounts to take in the evaluation of IO, the function:

$$W\left(F_\pi(.|c,e)\right) \;=\; \frac{\mu_c}{\mu}(1-G_c),$$

where μ_c/μ is the ratio between the average outcome associated to the distribution conditional on circumstance c and the population average, while G_c is the Gini coefficient of circumstance c's distribution.[20]

For every W, the IO index always allows to rank social states, although the conclusion on changes in equality of opportunity is not robust with respect to the evaluation of advantage. Yet, the index is consistent with an anonymous opportunity equalization criterion: if anonymous EZOP is satisfied, one should have $IO(0) \geq IO(1)$ for all preferences W.

3.3 Aggregation in the effort dimension

Let us now consider a situation where effort can be summarized by a scalar indicator $e \in \mathbb{R}^+$. We refer to the distribution of effort within a type by $G(e|c,\pi)$.

Consider the anonymous or non-anonymous equalization principles. Assume first that effort is realized and observable. This corresponds to what has been referred to in the EOP literature as an *ex post* situation.[21] A straightforward extension of definitions 2 and 3 to the multiple effort setting can be made by requiring equalization to hold at every effort level. With ideal data, the anonymous or non-anonymous equalization criteria can be implemented and separately tested at every effort level.

In most existing data sets, however, information on effort is missing. In this context, it is only possible to observe for each type its outcome distribution, given by:

$$F_\pi(y|c) \quad = \quad \int_E F_\pi(y|c,e)dG(e|c,\pi). \tag{3}$$

are ordered. Overall welfare depends on the evaluation of how much dispersed are the average realizations associated to every type.

[20] For a complete survey of Gini-type indices for Equality of Opportunity sets, see Weymark (2003).

[21] See for instance Fleurbaey and Peragine (2013)

In the presence of luck, the distribution of outcome of a given type arises from a *mixture of luck and effort* factors. Hence, contrary to Roemer (1998), it is not possible to identify effort with the quantiles of this distribution.

The ex-ante approach Although the distributions $F_\pi(.|c)$ do not allow to assess ex post equalization, they are interesting in their own right and relevant for opportunity equalization. Each distribution captures the opportunity sets associated to different types in an *ex ante* perspective, i.e. before the effort choices are made. If equalization judgements are made without knowing in advance what individual effort choices will be, the ex post level of effort could be treated as luck. This amounts to assume that all individuals in a type exert similar effort. One may further assume that effort levels are comparable across types, as discussed below. This comes close in spirit to the analysis of Van de gaer (1993). In this case, equalization should be decided on the basis of the outcome distributions of each type, $F_\pi(y|c)$. We refer to this criterion as ex ante equalization, defined by:

Definition 4 (Ex ante non-anonymous EZOP between multiple types) *Moving from state $\pi = 0$ to $\pi = 1$ equalizes opportunity ex ante over the set of circumstances C on the set of preferences $\mathcal{P}$ if and only if for all preferences $W \in \mathcal{P}$, for all $(i, j) \in \{1, \ldots, T\}$, we have:* $|\Delta_W(F_0(.|c_i), F_0(.|c_j))| \geq |\Delta_W(F_1(.|c_i), F_1(.|c_j))|$.

According to this definition, opportunities are equalized if every preference agrees that the gap between the expected opportunity sets associated to every pair of circumstances falls by effect of the change in social state. Here, opportunity sets are "expected" in the sense that they are evaluated before individuals make their effort choice.

When $\mathcal{P} = \mathcal{R}$, proposition 5 can be used to identify ex ante non-anonymous EZOP. The same approach can be used to identify the anonymous approach.

The Roemerian setting We now consider the special case of the Roemerian setting[22] where luck plays no role: individual outcomes only depend on circumstances and effort. Individuals with circumstances c and effort e in state π are assigned with a single value of outcome $Y_\pi(c, e)$. Since luck plays no role, ex post equalization amounts to require that for all (c, c') and all e: $|Y_0(c, e) - Y_0(c', e)| \geq |Y_1(c, e) - Y_1(c', e)|$.

[22]For a complete discussion of the conditions of identification of equality of opportunity in Roemer's model, see O'Neill, Sweetman and Van De Gaer (2000) and Lefranc et al. (2009)

Roemer's view of equality of opportunity further requires, on a priori grounds, that effort be defined in such a way that its distribution is independent of type. The argument is that since individuals cannot be held responsible for their type, they should not be held accountable for the association between their "effort" and their type. One may push the argument further and require that the distribution of effort be also independent of the state. In this case, we have that for all c and π, $G(e|c,\pi) = G(e)$. Furthermore, under the assumption that the outcome function $Y(c,e)$ is strictly increasing in e, the individual effort within a type can be identified by the rank in the type-specific outcome distribution: hence, an individual with outcome y and circumstances c in state π will have exerted effort $F_\pi(y|c)$.

In the Roemerian setting, the effort is normalized so that its distribution is uniform over the interval $[0,1]$, and $Y_\pi(c,e)$ is simply given by $F_\pi^{-1}(e|c)$. Thus requiring ex post EZOP in this setting amounts to require that, for all $p \in [0,1]$: $|F_0^{-1}(p|c) - F_0^{-1}(p|c')| \geq |F_1^{-1}(p|c) - F_1^{-1}(p|c')|$. This shows that the absolute gap curve dominance condition, defined in proposition 1, turns out to be a necessary and sufficient condition for ex post EZOP in the Roemerian setting. Furthermore, as a consequence of proposition 1, this condition is necessary for ex ante EZOP. As a result, ex ante EZOP implies ex post EZOP in the Roemerian setting.

The general case In the general case where luck and effort distributions are not degenerate, the relationship between ex ante and ex post equalization cannot be established without further assumptions. This can be illustrated by a simple example. Consider two circumstances, c and c', and many effort levels. Assume that for all effort levels, type c dominates c' at the first order. In this case, ex post EZOP requires that for all e, $|F_0(y|c,e) - F_0(y|c',e)| \geq |F_1(y|c,e) - F_1(y|c',e)|$. Assume further that effort is distributed independently of type and state. Under these two assumptions, we have, using (3): $|F_\pi(Y|c) - F_\pi(y|c')| = \int |F_\pi(y|c,e) - F_\pi(y|c',e)|dG(e)$. This allows to establish that ex post EZOP implies ex ante EZOP. However, this is only valid under the two maintained assumption. Unfortunately, these assumptions cannot be tested empirically, without observing effort.

This shows that in the general case, ex post equalization cannot be identified using ex ante comparisons. Furthermore, failing to accept ex ante equalization provides little

guidance on the fact that the ex post criteria also fails to be accepted, unless one is willing to make non-testable assumption of monotonicity of the ranking of circumstances with respect to effort.

4 Empirical application: Child care expansion and equalization of opportunity in Norway

Recently, policymakers both in the US and in Europe are pushing for expanding access to child care, in an effort to alleviate early life differences across socioeconomic groups. Indeed, early childhood investments are often seen as the means *par excellence* to equalize life chances (Blau and Currie 2006, Heckman and Masterov 2007). To illustrate the usefulness of our framework for policy evaluations, we now apply it to evaluate the long term impact of a large scale child care reform in Norway.

The Kindergarten Act passed the Norwegian parliament in June 1975. It assigned responsibility for child care to local municipalities and was followed by large increases in federal funding. The reform constituted a substantial positive shock to the supply of subsidized child care, which had been severely constrained by limited public funds. The child care coverage rate for 3 to 6 year olds increased from less than 10 % in 1975 to over 28 % by 1979.[23]

Our objective is to assess whether the expansion of child care equalized opportunity among Norwegian children. The outcome variable we focus on is individual earnings at age 30–36. Our circumstance variable is parental earnings during early childhood. Havnes and Mogstad (2011) show that the child care expansion had, on average, positive long-run effects on children's education and labour market attachment. Havnes and Mogstad (2014) document that the effects were highly heterogenous with gains clustered in the lower end of the earnings distribution, and on children from disadvantaged backgrounds.

We extend on Havnes and Mogstad (2014) by looking at the full distributional consequences of child care expansion within family background, and by bringing the EZOP framework to bear on these results. Specifically, we examine to what extent the expansion of child care equalized children's earnings distributions as adults, conditional on parental earnings deciles.

[23]For detailed information about the program and for descriptive statistics, we refer the reader to Havnes and Mogstad (2011).

4.1 Empirical implementation

Assessing whether the Kindergarten Act equalized opportunities across Norwegian children requires two sets of outcome distributions: The distributions of observed outcomes by family background among children who have experienced the childcare expansion, and the counterfactual distributions that would have prevailed in absence of the reform. Following Havnes and Mogstad (2014), we apply a difference-in-difference (DiD) approach, exploiting the fact that the supply shocks to subsidized child care were larger in some areas than in others. Specifically, we compare the adult earnings of children aged 3 to 6 years old before and after the reform, from municipalities where child care expanded a lot (i.e. the treatment group) and municipalities with little or no increase in child care coverage (i.e. the comparison group).

We focus on the early expansion, which likely reflects the abrupt slackening of constraints on the supply side caused by the reform, rather than a spike in the local demand. We consider the period 1976–1979 as the child care expansion period. To define the treatment and comparison group, we order municipalities according to the percentage point increase in child care coverage rates over the expansion period. We then separate the sample at the median, the upper half constituting the treatment municipalities and the lower half the comparison municipalities. To define the pre-reform and post-reform groups, we exploit that children born 1967–69 enter school before the expansion period starts, while children born 1973–76 are in child care age after the expansion period has ended. Havnes and Mogstad (2011, 2014) show that the expansion of child care is not easily explained by observable characteristics.

To estimate the impact of the reform on the distribution of children's earnings, conditional on parental earnings, we estimate the following equation:

$$\mathbb{1}\{y_{it} \geq y\} = \gamma_t(y) + [\boldsymbol{\beta_0}(y) + \boldsymbol{\beta_1}(y)P_t + \boldsymbol{\beta_2}(y)T_i + \boldsymbol{\beta_3}(y)T_i \cdot P_t] \cdot \mathbf{x}_{it} + \epsilon_{it}(y) \qquad (4)$$

where $\mathbb{1}\{\cdot\}$ is the indicator function, y_{it} are average earnings in 2006–2009 of child i born in year t, and y is a threshold value of earnings discussed below. T_i is a dummy equal to one if the child is from a treatment municipality and zero otherwise, and P_t is a dummy equal to one for post-reform cohorts and zero otherwise. γ_t is a birth cohort fixed effect, and ϵ is the error term. The vector $\mathbf{x}_{it}$ contains a fourth-order polynomial in the

average yearly earnings of the child's parents when the child was in child care age, that is $\mathbf{x}'_{it} = \left(x_{it}, x_{it}^2, x_{it}^3, x_{it}^4\right)$.[24]

The vector $\boldsymbol{\beta_3}(y)$ has dimension (4×1), and provides DID-estimates of how the reform affected the earnings distribution of affected children. In the spirit of standard DiD, the estimator uses the observed change *in the distribution* around the value y, from before to after treatment, as an estimate of the change that would have occurred in the treatment group over this period in the absence of treatment. This change is measured by the effect of the covariates on the recentered influence function (RIF) of the income distribution of interest at income levels y.[25] The identifying assumption underlying the RIF-DiD estimator is that *the change in population shares* from before to after treatment around a given *level of earnings* would be the same in the treatment group as in the comparison group, in the absence of the treatment.

Note that equation (4) allows for heterogeneity in the effect of the reform on the distribution of earnings along two dimensions. First, $\boldsymbol{\beta_3}(y)$ is a function of the threshold earnings so the effect of the reform is allowed to vary along the earnings distribution. Second, since $\boldsymbol{\beta_3}(y)$ is interacted with a polynomial in parental earnings ($\mathbf{x}_{it}$), the effect of the reform is allowed to vary according to family background.

Equation (4) provides estimates defined in terms of changes in probability mass at each value y. From these, we can compute the change in earnings induced by the reform by rescaling with an estimate of the density at y (Firpo et al. 2009a). When y is a quantile, this yields an estimate of the quantile treatment effect (QTE).

Our EZOP criterion rests on a comparison of the effects of the reform at quantiles of the earnings distribution conditional on circumstances. For each circumstance c and each quantile $p \in [0, 1]$, define $Q_1(p|c) = F(y|c, T = 1, P = 1)$ as the value of the p^{th} quantile in the actual distribution of earnings among treated children, conditional on circumstances. The estimated QTE at quantile p for children with circumstances c can then be defined as:

$$QTE(p|c) = \frac{\mathrm{E}\left[\boldsymbol{\beta_3}(Q_1(p|c)) \cdot \mathbf{x}_{it} | C_{it} = c\right]}{f\left(Q_1(p|c) | C_{it} = c\right)} \tag{5}$$

[24] We tested alternative polynomial specifications without any impact on our results.

[25] See Firpo, Fortin and Lemieux (2009a) for a discussion of the RIF-estimator, and Athey and Imbens (2006a) or Havnes and Mogstad (2014) for a discussion of non-linear difference-in-differences methods.

where C_{it} denotes the circumstances of individual i born in cohort t, and $f(\cdot|\cdot)$ denotes the density of the earnings distribution $F(\cdot|\cdot)$. Because $QTE(p|c)$ estimates the impact of the treatment, we readily construct an estimate of the counterfactual quantile in the absence of treatment as $Q_0(p|c) = Q_1(p|c) - QTE(p|c)$.

In the empirical application, we use the earnings decile of the child's parents to define circumstances, and estimate equation (4) using OLS at each percentile of the earnings distribution conditional on circumstances. We then use a kernel estimate of the density from this distribution to construct our estimate of $QTE(p|c)$. Our estimation sample is based on Norwegian registry data, and covers the population of Norwegian children born to married mothers in the relevant cohorts, as discussed in Havnes and Mogstad (2014). Standard errors are obtained using a non-parametric bootstrap with 300 replications. Based on our estimates of the actual and counterfactual outcome distributions and on the bootstrapped covariances, we implement stochastic dominance tests as discussed in appendix B.

4.2 Results for three classes

To evaluate how the child care expansion affected inequality of opportunity, we now apply our EZOP framework, defining children's circumstances from parental earnings deciles. This involves a large number of pairwise comparison, to which we return below. To clarify the intuition behind the comparisons, and to illustrate graphically the implementation of our framework and the effect of the reform, however, we first focus on three types in the population: Children whose parents had earnings in the second, the fifth and the ninth decile, respectively. For expositional simplicity, we will usually refer to these simply as lower class, middle class and upper class children.

We start by analyzing the extent of inequality of opportunity before the implementation of the child care expansion. Panel (a) in figure 1 presents the counterfactual distributions $Q_0(p|c)$ that would have been observed in the absence of the policy ($\pi = 0$). The figure shows first order stochastic dominance when we compare any pair of distributions. This indicates that equality of opportunity is clearly violated. Furthermore, for all preferences under risk, there is a clear ordering of family types, with upper class children doing better than middle class children, and middle class children doing better than lower class children.

Panel (b) in figure 1 shows the impact of the child care expansion on the earnings distribution of children in these three groups. The dashed line presents the QTE for middle

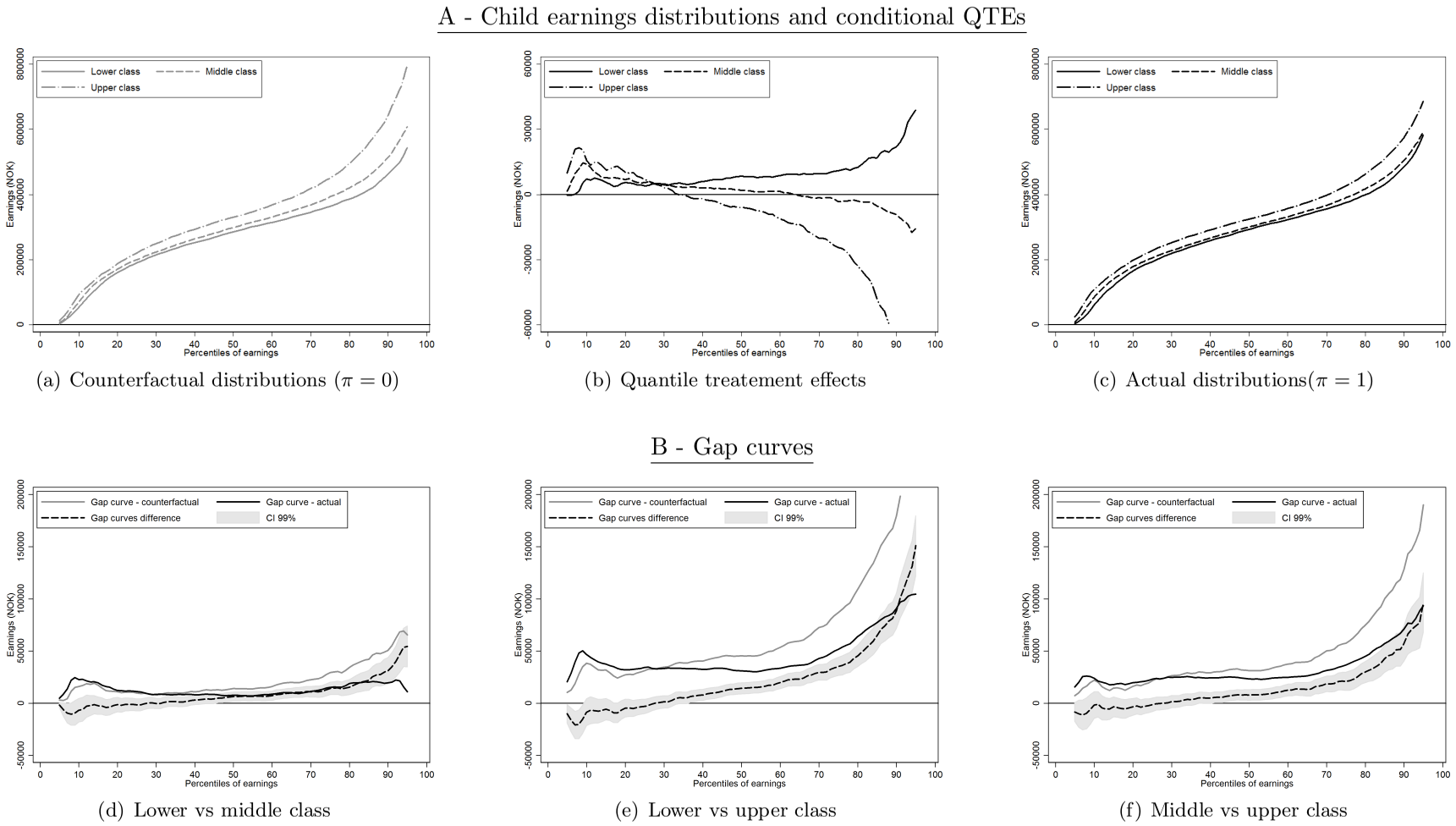
class children. Overall, the effect of the child care expansion in this group are relatively modest. However, there is significant heterogeneity in the impact of the policy: *within* the middle class, effects are positive for children sitting in the bottom of the earnings distribution, and turn negative in the upper end of the distribution. The dotted line gives the effect on upper class children. In this group, the reform has a modest positive impact for children in the bottom of the conditional distribution but had large and negative impact in the top of the distribution. Lastly, the solid line provides estimates of the effect of the child care expansion for lower class children. On average, lower class children seem to benefit more from the child care expansion than children from middle and upper class. Furthermore, the heterogeneity in the effect of child care stands in marked contrast with what was observed in the other two groups: Among lower class children, the reform had a small positive effect in the bottom of the distribution but had increasingly large and positive effect as we move up the conditional earnings distribution.This suggests two likely conclusions. First, on average child care appears substitutable to parental resources, captured here by the class of origin. Second, the impact of the child care seems complementary to child's idiosyncratic resources, within the lower class, while the opposite seems to hold in the middle and upper class.

Panel (c) of figure 1 presents the conditional distribution of earnings after the policy implementation ($\pi = 1$). The figure shows first order stochastic dominance when we compare any pair of distributions. Hence, equality of opportunity does not prevail, even after the implementation of the reform. However, compared to panel (a), the gap between any pair of curves seems to have fallen at almost every quantile of the earnings distribution, which suggests that the implementation of the child care policy might have partially equalized opportunities across the three classes.

To implement our EZOP procedure, we present in panels (d)–(f) the estimated gap curves from pairwise comparisons of children from different family types under both social states, alongside gap curve differences between these states with a 99% confidence interval bands.[26] Since the conditional distribution can be ordered according to first-order stochastic dominance, we may invoke proposition 2: Gap curve dominance provides a necessary and sufficient condition for equalization of opportunity.

Two main features stand out. First, in both social states, gap curves are virtually

[26]Remember from section 2.2 that gap curves are constructed by taking differences between inverse cdf's at each percentile.



(a) Counterfactual distributions ($\pi = 0$)

(b) Quantile treatement effects

(c) Actual distributions($\pi = 1$)

(d) Lower vs middle class

(e) Lower vs upper class

(f) Middle vs upper class

Figure 1: Child earnings distributions, quantile treatment effects and gap curves

Note: The lower, middle and upper classes refer to selected family labor income deciles. Gap curves are defined as the vertical distance between the conditional cdfs of the different classes. Since the cdfs are ordered, the gap curves are always positive. Gap curves differences correspond to the vertical gap between the gap curves in the actual and the counterfactual states. Confidence intervals for the difference in gap curves at every conditional quantile are bootstrapped.

always positive. This reflects the stochastic dominance between all groups both with and without the child care reform. Second, the actual gap curve ($\pi = 1$) is almost always below the counterfactual gap curve ($\pi = 0$). This indicates that the reform reduced inequality of opportunity between all pairs of types. This fact is clarified by looking at the gap curve differences: While the difference is small and not statistically significant at the bottom of the distribution, the difference becomes positive and strongly statistically significant as we move up in the distribution.

The graphs in figure 1 are based on estimated quantiles. An informed conclusion about whether the child care reform led to equalization of opportunity should rest on a joint test of stochastic dominance in each pairwise comparison (i) for the actual distribution, (ii) for the counterfactual distributions and (iii) for the gap curves. Results of these tests are presented in Table 1. Panel A and B present test statistics from the counterfactual and actual setting, respectively. Note that the bootstrapped covariances are used to account for the QTE estimation error. We perform three tests. First, the null hypothesis is that the two distributions are equal. This is strongly rejected in all comparisons. Second, the null hypothesis is that the distribution of the underprivileged group first order stochastic dominates the distribution of the privileged group, i.e. lower class over middle class over upper class. Not surprisingly, the hypothesis is strongly rejected in all comparisons.[27] Third, we test for the reverse relation: That the underprivileged are dominated by the privileged. In this case, the hypothesis cannot be rejected in any of the comparison, with p-values always above 0.94.

Finally, panel C presents the main tests of equalization of opportunity. First, the null hypothesis is that the reform had no impact on inequality of opportunity (*neutrality*). This hypothesis is strongly rejected by the data. Second, the null hypothesis is that the reform equalized outcomes across children from different classes (*equalization*). This hypothesis cannot be rejected by the data, with p-values above 0.38. Third, the null hypothesis is that the reform disequalized outcomes, increasing inequality of opportunity. In this case, we can again strongly reject the hypothesis in all comparisons.

To summarize, the analysis shows first that the ordering between children from different classes in terms of their labor market performance is quite clear in Norway: Upper class

[27]Both in panels A and B of table 1, the values of the tests statistics taken under the null hypothesis of equality and dominance coincide. This result is a consequence of the definition of the test statistics presented in the appendix.

Table 1: Joint dominance and equality tests for actual and counterfactual children earning distributions and gap curves, for selected parental earnings deciles: Wald-test statistics and associated p-values (in brackets) for various null hypothesis

	Pairwise comparisons of social classes:					
	Lower vs middle		Lower vs upper		Middle vs upper	
A - Cdfs, counterfactual setting ($\pi = 0$)						
$H_0 : \sim$	72.9	[0.000]	659.4	[0.000]	384.2	[0.000]
$H_0 : \succcurlyeq$	72.9	[0.000]	659.4	[0.000]	384.2	[0.000]
$H_0 : \preccurlyeq$	0.0	[0.944]	0.0	[0.949]	0.0	[0.947]
B - Cdfs, actual setting ($\pi = 1$)						
$H_0 : \sim$	40.1	[0.003]	423.7	[0.000]	266.3	[0.000]
$H_0 : \succcurlyeq$	40.1	[0.000]	423.7	[0.000]	266.3	[0.000]
$H_0 : \preccurlyeq$	0.0	[0.949]	0.0	[0.952]	0.0	[0.948]
C - Gap curves ($\pi = 0$ **vs** $\pi = 1$)						
H_0 : Neutrality	84.2	[0.000]	266.4	[0.000]	125.0	[0.000]
H_0 : Equalization	4.8	[0.672]	11.2	[0.381]	9.1	[0.468]
H_0 : Disequalization	76.0	[0.000]	248.4	[0.000]	112.0	[0.000]

Note: The lower, middle and upper classes refer to selected parental earnings deciles groups. In panels A and B, for each of the three pairs of classes, we test the following three null hypothesis: equality of the cdfs ($\sim$), first-order stochastic dominance of the worse-off class over the well-off class ($\succcurlyeq$) and first-order stochastic dominance of the well-off class over the worse-off class ($\preccurlyeq$). In panel C, for each pair of classes, we compare gap curves under the actual and counterfactual states and test three null hypothesis: the gap curves are statistically equal (neutrality); the gap curve in the counterfactual policy state is everywhere larger than in the actual policy state (equalization); the gap curve in the counterfactual policy state is everywhere smaller than in the actual policy regime (disequalization). Gap curves are defined according to the order of groups estimated from panels A and B. Covariances are bootstrapped. Joint tests are performed on ventiles of child earnings distributions.

children do better than middle class children, and middle class children do better than lower class children. Second, the analysis shows that the child care reform in 1975 did indeed equalize substantially the opportunities across children from different classes. Using the Gini-type evaluation function, we can also assess how the policy affected the value of the opportunity sets of each type. For the bottom two classes, results indicate that the reform had a positive effect: The opportunities of the lower and middle class children increased by 4.3% and 3% respectively. In contrast, the value of the opportunity set of the upper class only increased by a modest 1%, which turns out to be statistically insignificant. This differential in growth rates indicates that the lower and middle classes benefited from the policy reform, both in absolute terms, and in relative terms, in the sense that they caught up with the upper class.[28] Third, the QTE estimates show that this equalization came both from positive impacts at the lower end of the distribution and from negative impacts at the upper end for many children. This raises a concern about the universal design of

[28]We bootstrapped 300 times the differences across policy states of the difference in Gini evaluation function attributed to lower, middle and upper classes. These differences (t-statistics) between lower and middle classes are 0.0073 (3.091), between lower and upper classes are 0.0206 (3.185) and between middle and upper classes are 0.0133 (2.693).

the child care expansion, as discussed in Havnes and Mogstad (2014).

4.3 Results for all parental earnings deciles

We now extend the analysis to consider children in groups based on each decile of the parental earnings distribution. The analysis follows the same format as above: For each group of children, we estimate QTE's, actual cdf's and counterfactual cdf's, and test for stochastic dominance and gap curve dominance. For brevity, we omit the specific results for each group. The final results from the dominance tests are summarized graphically in figure 2. Above the diagonal, we test dominance of column groups over row groups, i.e. of the disadvantaged group over the advantaged group. Below the diagonal, we test instead dominance of the advantaged group over the disadvantaged group. We use the shading of the block to indicate the p-value on rejection of the null hypothesis of dominance, with darker blocks indicating higher p-values. Dark blocks indicate failure to reject the null hypothesis of dominance, while light blocks indicate that the null hypothesis is rejected. Finally, we also test for equality of the earnings distributions across types, and indicate failure to reject equality with a black bullet inside the block.

To clarify, compare to the analysis with the three classes above. The squares in row 2, column 9, and in row 9, column 2 compare children from lower class (D2) to children from upper class (D9). This essentially summarizes the findings from the second column of table 1. Consider Panel A of figure 2, where we test for dominance in the actual setting. The dark color in row 9, column 2 indicates the failure to reject dominance in the third line of panel A in table 1. Similarly, the light color in row 2, column 9 indicates the rejection of dominance in the third line of panel A in table 1. The absence of bullets in both squares indicates that we reject equality of the earnings distributions, as in the first line of panel A in table 1.

Return now to the overall evaluation of whether the child care expansion equalized opportunity across family types. Panels A and B of figure 2 report the results of first-order stochastic dominance tests in the actual and the counterfactual state. In both states, the results suggest a strong monotonic relation between parental earnings and the earnings advantage of children. Above the diagonal, we universally fail to reject the hypothesis that the earnings distribution of more advantaged children dominates that of less advantaged children. In contrast, below the diagonal, we can reject that the earnings distribution of less

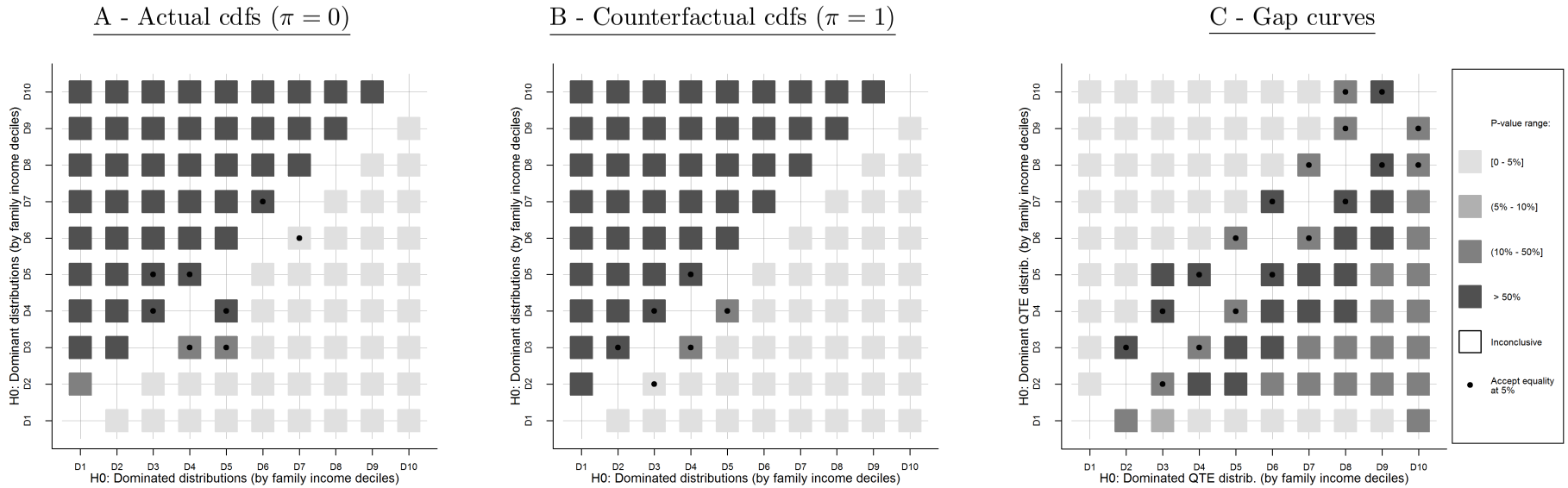


Figure 2: Joint dominance and equality tests for actual and counterfactual children earning distributions and gap curves, conditional on parental earnings decile

Note: The groups D1–D10 refer to parental earnings deciles. Panels A and B reports p-values of joint tests of first-order stochastic dominance and equality of cdfs for pairs of groups, in the counterfactual and actual regime. Each square in the matrix compares the distribution of the type on the horizontal axis with the type on the vertical axis. Squares below the 45° line correspond to tests of the null hypothesis that the type on vertical axis is dominated by the type on the horizontal axis. Squares above the 45° line correspond to tests of the null hypothesis that the type on vertical axis dominates the type on the horizontal axis. Panel C reports p-value of joint tests of dominance and equality for gap curves performed through QTE comparisons across groups (see appendix B.5 for details). Squares below the 45° line correspond to tests of the null hypothesis that the gap curve under $\pi = 1$ is below the gap curve under $\pi = 0$, indicating that the disadvantaged group gains more from the policy than the advantaged group. Squares above the 45° line correspond to tests of the opposite hypothesis. The color of the square indicates the p-value for the rejection of the null hypothesis. Black squares indicate rejection of the null. Dots in the figure indicate when equality between the distribution of the two types cannot be rejected at conventional level. Joint dominance tests are based on children earnings ventiles. Covariances are bootstrapped.

advantaged children dominates that of more advantaged children virtually everywhere. The only exceptions are three central comparisons around the diagonal, where the differences in parental earnings across groups is rather small, and equality cannot be rejected and/or we reject dominance in both directions. Overall, these tests provide clear evidence of inequality of opportunity for earnings among Norwegian children in both states. Moreover, the ordering of groups is highly stable across policy regimes, so that anonymous and non-anonymous EZOP comparisons are operationalized by the same tests.

We now turn to the test of equalization of opportunity. Panel C of figure 2 reports the results of gap curve dominance tests for all pairs. In most cases, the estimates suggest that the implementation of the policy significantly decreases the opportunity gap between the advantaged and the disadvantaged type. Below the diagonal, we almost never reject an improvement in the position of the less advantaged children compared to more advantaged children. Above the diagonal, we almost always reject an improvement in the position of the more advantaged children compared to less advantaged children. In other words, at all ranks, children raised in less advantaged families seem to benefit more (or lose less) from the child care expansion than do children raised in more advantaged families.

Overall this equalization of opportunity can be quantified by resorting to a specific inequality of opportunity index. Using the Gini Opportunity index of Lefranc et al. (2008) previously defined, we find that unfair inequality decreased by 8.8% as a result of the expansion in kindergarten provision.[29]

There are however, some exceptions to this equalization of opportunity. These exceptions are again mostly along the diagonal, where parental earnings differences are less pronounced. Here, we usually fail to reject equality between gap curves, and fail to reject equalization in either direction. Another important exception is the group of children with parental earnings in the first decile. For these children, we reject equalization when we compare to children from deciles 4–9 in the parental earnings distribution. This suggests that the child care reform partly failed to equalize opportunity for the most disadvantaged group.

The analysis above shows clearly the change in the pattern between parental earnings and children's advantage following the child care expansion. It does not, however, discriminate between a homogeneous narrowing of the gap across all ranks of the children's

[29]The Gini Opportunity indices (standard error) are $GO(0) = 0.0358$ (0.0013) and $GO(1) = 0.0326$ (0.001). The p-value for $H_0 : GO(0) = GO(1)$, based on bootstrapped standard errors, is 0.029.

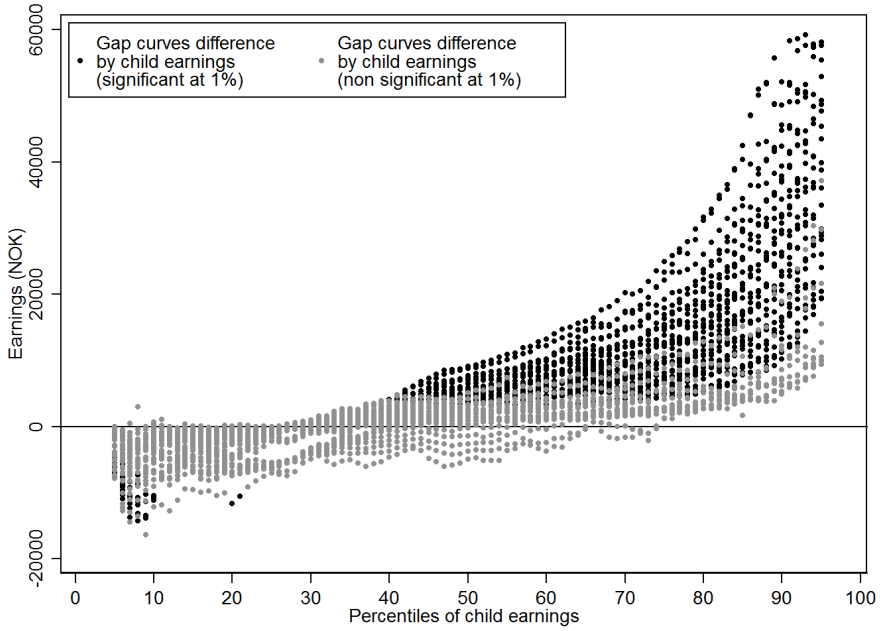


Figure 3: Equalization of opportunity across family labor income deciles types.

Note: The figure reports, for every percentile of the sons earnings distribution, the differences in QTE associated to pairs of different groups, i.e. the difference between the gap curves of the two types. Disjoint tests of equality of the QTE are performed using bootstrapped standard errors. Differences in QTE that are not significantly different form zero are reported in gray. Groups are ordered according to ISD at order one. Out of 10 groups of parental earnings, 45 pairs of comparisons are performed at every percentiles of the sons earnings distribution.

earnings distribution, and an overall narrowing where some parts of the distribution contribute more than others. To clarify this point, figure 3 reports the difference between gap curves in the actual and counterfactual state, for each percentile of the children's earnings distribution. At each percentile of the conditional earnings distributions of the children, we report the difference between the gap curves at that percentile from each of the 45 comparisons below the diagonal in panel C of figure 2. We restrict to these 45 comparisons since the order of the groups is completely determined in panel A and B of figure 2. In the figure, we indicate by a dark color the gap curve differences at a given percentile that are statistically significant at the 1% level.

Figure 3 shows that equalization among most of the groups is driven not by a reduction in the gap between children from poor families, but rather by a convergence in the earnings distributions between the median and the top of the distribution. This is driven partly by the fact that estimated effects are rather homogenous across groups at the lower end of the distribution, and partly by the fact that the negative QTE-estimates at the upper end of the distribution are particularly large for advantaged groups.

5 Concluding remarks

The first contribution of this paper is theoretical. We develop a new criterion for ranking social states from the equality of opportunity perspective. Our criterion for *equalization of opportunity* entails a difference-in-differences comparison of outcome distributions conditional on circumstance. First, types are compared *within* each social state separately, to assess the direction and distribution of unfair advantage across all possible pairs of types. In cases where ordering of types is not unanimous, we proceed by minimally refining the set of potential preferences until agreement is reached. We show that this refinement is easily implementable using inverse stochastic dominance tools.

Second, differences are taken *between* social states in order to assess *changes* in the extent and distribution of unfair advantage. We propose an innovative model based on comparisons of changes in the distance between pairs of distributions. Our criterion requires unanimity, within a large class of preferences, in the evaluation of the fall in the illegitimate advantage enjoyed by one type with respect to another. We study identification procedures and implementation issues, showing the equivalence of our EZOP order with gap curve dominance.

Our results extend to the equality of opportunity framework some important results in social welfare ordering. Several authors have investigated the relationship between stochastic dominance orders and social orders in a welfarist context. A well known result in this literature is that all inequality averse social welfare function give a unanimous ranking of two social states if and only if the distributions of these social states can be ranked according to second-order stochastic dominance (Kolm 1969, Atkinson 1970, Shorrocks 1983). Aaberge et al. (2014) generalize this result and show the equivalence of higher order inverse stochastic dominance and nested subclasses of social welfare functions that are axiomatically justified and have appealing interpretation. Our analysis builds on this approach but introduces an important distinction. Instead of focusing on inequality of outcomes, as is the case in the welfarist approach, our social order criterion is based on the principle of equality of opportunity, in line with modern theories of distributive justice. Our results indicate that the equivalence between unanimous welfarist ordering and stochastic dominance ordering of outcome distributions can be transposed to the equality of opportunity setting if we focus on the opportunity gap curve between types.

Our second contribution is to provide a statistical framework that allows to implement our equalization of opportunity criterion. In particular, we provide inference procedures to test whether the equalization conditions are empirically supported by the data. Our application also underlines that econometric models allowing for heterogenous effects can be tightly connected to the normative assessment of distributional issues. The recent econometric literature has provided important tools for estimating the heterogenous impact of policy intervention on some outcome of interest.[30] Since our equalization criterion can be expressed in terms of restictions on quantile treatment effects, this paper suggests a simple way in which these estimates can be used to assess whether the heterogenous effect of a given policy help promote distributive justice.

The third contribution of this paper pertains to the empirical analysis of the effectiveness of early childhood intervention at equalizing life chances. Growing evidence on the role of family background on lifelong earnings potential (Björklund and Salvanes 2011, Black and Devereux 2010) has brought educational policies to the forefront as potential tools for alleviating differences stemming from family background. This has taken particular prominence due to theory and evidence suggesting that skill formation early in life may be crucial in determining children's trajectories (Cunha and Heckman 2007). Expanding access to quality child care may be expected to equalize opportunities among treated children, by weakening the dependence between family background and children's development. The literature on the effect early childhood investments in general, and child care in particular, provides mixed conclusions on this issue.[31] While studies of targeted programs often find positive effects (see e.g. Barnett (1995), Currie (2001), and Karoly, Kilburn and Cannon (2005) for surveys), the literature on universal programs is smaller and findings are mixed.[32] We extend this literature by providing evidence on the impact of a universally available large scale child care program on long run equality of opportunity.

Applying our framework to reevaluate the introduction of universally available child

[30] Abadie, Angrist and Imbens (2002), Chernozhukov and Hansen (2005), Athey and Imbens (2006b), Firpo (2007) and Firpo, Fortin and Lemieux (2009b) are important contributions to this literature. The RIF-DID estimator of Havnes and Mogstad (2014) which we apply in this paper belongs to the same econometric vein.

[31] For recent reviews, see Almond and Currie (2011), Ruhm and Waldfogel (2012), or Baker (2011).

[32] Several studies from Canada find negative effects (Baker, Gruber and Milligan 2008, Lefebvre, Merrigan and Verstraete 2008, DeCicca and Smith 2013), while zero impact has been found in the US and Denmark (Cascio 2009, Gupta and Simonsen 2010). Meanwhile, positive impacts on long run outcomes are found in the US (Fitzpatrick 2008), Uruguay (Berlinski, Galiani and Manacorda 2008), Norway (Havnes and Mogstad 2011, Havnes and Mogstad 2014), Germany (Dustmann, Raute and Schønberg 2013, Felfe and Lalive 2013), and Spain (Felfe, Nollenberger and Rodriguez-Planas 2013).

care in Norway, we conclude that kindergarten expansion indeed equalizes opportunities among children. In our empirical analysis, we expand on previous evidence to emphasize the highly heterogenous effects on children's adult earnings across child family backgrounds, as defined by parental earnings. We document that impacts were in fact quite homogenous across family backgrounds in the lower end of the distribution, with positive effects, but highly dispersed in the upper end, where disadvantaged gained while advantaged lost from the expansion. When applied to our EZOP framework, the results suggest that the child care expansion significantly equalized opportunities between children from most family backgrounds. Our results emphasize that early intervention may be an important tool to promote equal life chances.

Two important caveats should be noted. First, echoing results in Havnes and Mogstad (2014), our results show that the equalization of opportunity resulting from the reform is driven importantly by reduced earnings at the upper end of the earnings distribution for affected children. While reduced earnings in the top of the distribution improves equality of opportunity, it does so in a somewhat displeasing way. An important question is whether resources devoted to provide child care for children from upper class families could be reallocated to improve quality of care or uptake for lower class children.

Second, although there is strong agreement on equalization of opportunity for the vast majority of groups, it is not possible to conclude in favor of equalization overall. In fact, the policy seems to increase the gap for the least succesful children in the most disadvantaged group compared to most other types. This result indicates that the Kindergarten Act produced relatively low returns for these children, leaving them even further behind compared to the children from somewhat less disadvantaged backgrounds, that benefitted handsomely. This finding casts a shadow on the effectiveness of large-scale child care for the neediest children, and deserves further investigation.

References

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

Aaberge, R., Havnes, T. and Mogstad, M. (2014). A theory for ranking distribution functions, *Memorandum 20/2014*, Department of Economics, University of Oslo.

Aaberge, R. and Mogstad, M. (2011). Robust inequality comparisons, *The Journal of Economic Inequality* **9**(3): 353–371.

Aaberge, R., Mogstad, M. and Peragine, V. (2011). Measuring long-term inequality of opportunity, *Journal of Public Economics* **95**(3-4): 193 – 204.

Abadie, A., Angrist, J. and Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings, *Econometrica* **70**(1): 91–117.

Almond, D. and Currie, J. (2011). *Human Capital Development before Age Five*, Vol. 4 of *Handbook of Labor Economics*, Elsevier, chapter 15, pp. 1315–1486.

Almås, I., Cappelen, A. W., Lind, J. T., Sørensen, E. . and Tungodden, B. (2011). Measuring unfair (in)equality, *Journal of Public Economics* **95**(7-8): 488–499.

Andreoli, F. (2013). Inference for inverse stochastic dominance, *ECINEQ working paper 295*.

Athey, S. and Imbens, G. W. (2006a). Identification and inference in nonlinear difference-in-differences models, *Econometrica* **74**(2): pp. 431–497.

Athey, S. and Imbens, G. W. (2006b). Identification and inference in nonlinear difference-in-differences models, *Econometrica* **74**(2): 431–497.

Atkinson, T. B. (1970). On the measurement of inequality, *Journal of Economic Theory* **2**: 244–263.

Baker, M. (2011). Innis lecture: Universal early childhood interventions: what is the evidence base?, *Canadian Journal of Economics* **44**(4): 1069–1105.

Baker, M., Gruber, J. and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being, *Journal of Political Economy* **116**(4): 709–745.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes, *The Future of Children* **5**(3): 25–50.

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

Berlinski, S., Galiani, S. and Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles, *Journal of Public Economics* **92**(5-6): 1416–1440.

Björklund, A., Jäntti, M. and Roemer, J. (2011). Equality of opportunity and the distribution of long-run income in sweden, *Social Choice and Welfare* pp. 1–22. 10.1007/s00355-011-0609-3.

Björklund, A. and Salvanes, K. G. (2011). *Education and Family Background: Mechanisms and Policies*, Vol. 3 of *Handbook of the Economics of Education*, Elsevier, chapter 3, pp. 201–247.

Black, S. E. and Devereux, P. J. (2010). Recent developments in intergenerational mobility, *in* D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 4B, North-Holland, Amsterdam, chapter 8.

Blau, D. and Currie, J. (2006). *Pre-School, Day Care, and After-School Care: Who's Minding the Kids?*, Vol. 2 of *Handbook of the Economics of Education*, Elsevier, chapter 20, pp. 1163–1278.

Bourguignon, F., Ferreira, F. H. G. and Menendez, M. (2007). Inequality of opportunity in Brasil, *Review of Income and Wealth* **53**(4): 585–618.

Cascio, E. U. (2009). Do investments in universal early education pay off? long-term effects of introducing kindergartens into public schools, *NBER Working Papers 14951*, National Bureau of Economic Research, Inc.

Chakravarty, S. R. and Dutta, B. (1987). A note on measures of distance between imcome distributions, *Journal of Economic Theory* **41**(1): 185 – 188.

Checchi, D. and Peragine, V. (2010). Inequality of opportunity in Italy, *Journal of Economic Inequality* **8**: 429–450.

Chernozhukov, V. and Hansen, C. (2005). An IV Model of Quantile Treatment Effects, *Econometrica* **73**(1): 245–261.

Cunha, F. and Heckman, J. (2007). The technology of skill formation, *American Economic Review* **97**(2): 31–47.

Currie, J. (2001). Early childhood education programs, *Journal of Economic Perspectives* **15**(2): 213–238.

Dardanoni, V., Fields, G., Roemer, J. and Sanchez Puerta, M. (2005). How demanding should equality of opportunity be, and how much have we achieved?, *in* S. L. Morgan, D. Grusky and G. Fields (eds), *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, Stanford University Press, Stanford.

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

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

DeCicca, P. and Smith, J. (2013). The long-run impacts of early childhood education: Evidence from a failed policy experiment, *Economics of Education Review* **36**: 41–59.

Dustmann, C., Raute, A. and Schønberg, U. (2013). Does universal childcare matter? evidence from a large expansion in pre-school education, *Working paper*, University College of London.

Dworkin, R. (1981). What is equality? Part 1: Equality of welfare, *Philosophy & Public Affairs* **10**(3): pp. 185–246.

Ebert, U. (1984). Measures of distance between income distributions, *Journal of Economic Theory* **32**(2): 266 – 274.

Felfe, C. and Lalive, R. (2013). Early child care and child development: For whom it works and why, *CEPR Discussion Papers 9274*, C.E.P.R. Discussion Papers.

Felfe, C., Nollenberger, N. and Rodriguez-Planas, N. (2013). Can't buy mommy's love? universal childcare and children's long-term cognitive development, *CESifo Working Paper Series 4069*, CESifo Group Munich.

Ferreira, F. H. G. and Gignoux, J. (2011). The measurement of inequality of opportunity: Theory and an application to latin america, *Review of Income and Wealth* **57**(4): 622–657.

Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects, *Econometrica* **75**(1): 259–276.

Firpo, S., Fortin, N. M. and Lemieux, T. (2009a). Unconditional quantile regressions, *Econometrica* **77**(3): 953–973.

Firpo, S., Fortin, N. M. and Lemieux, T. (2009b). Unconditional quantile regressions, *Econometrica* **77**: 953–973.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement, *B.E. Journal of Economic Analysis and Policy* **8**(1): 1–38.

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

Fleurbaey, M. and Peragine, V. (2013). Ex ante versus ex post equality of opportunity, *Economica* **80**(317): 118–130.

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

Gupta, N. D. and Simonsen, M. (2010). Non-cognitive child outcomes and universal high quality child care, *Journal of Public Economics* **94**(1-2): 30 – 43.

Havnes, T. and Mogstad, M. (2011). No child left behind: Subsidized child care and children's long-run outcomes, *American Economic Journal: Economic Policy* **3**: 97–129.

Havnes, T. and Mogstad, M. (2014). Is universal child care leveling the playing field, *Journal of Public Economics* **forthcoming**.

Heckman, J. J. and Masterov, D. V. (2007). The productivity argument for investing in young children, *Working Paper 13016*, National Bureau of Economic Research.

Karoly, L. A., Kilburn, M. R. and Cannon, J. S. (2005). *Early Childhood Interventions: Proven Results, Future Promise*, RAND Corporation, Santa Monica, CA.

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

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

Lefebvre, P., Merrigan, P. and Verstraete, M. (2008). Childcare policy and cognitive outcomes of children: Results from a large scale quasi-experiment on universal childcare in canada, *Cahiers de recherche 0823*, CIRPEE.

Lefranc, A., Pistolesi, N. and Trannoy, A. (2008). Inequality of opportunities vs. inequality of outcomes: Are western societies all alike?, *Review of Income and Wealth* **54**(4): 513–546.

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

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

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

O'Neill, D., Sweetman, O. and Van De Gaer, D. (2000). Equiality of opportunity and kernel density estimation: An application to intergenerational mobility, *Advances in Econometrics : A research Annual* **14**: 259–274.

Peragine, V. (2002). Opportunity egalitarianism and income inequality, *Mathematical Social Sciences* **44**(1): 45 – 64.

Peragine, V. (2004). Ranking income distributions according to equality of opportunity, *Journal of Economic Inequality* **2**: 11–30.

Ramos, X. and Van de gaer, D. (2012). Empirical approaches to inequality of opportunity: Principles, measures, and evidence, *Technical report*, ECINEQ WP 2012 – 259.

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

Roemer, J. and Trannoy, A. (2014). *Equality of opportunity*, Vol. Handbook of Income Distribution.

Ruhm, C. J. and Waldfogel, J. (2012). Long-term effects of early childhood care and education, *Forthcoming in Nordic Economic Policy Review* .

Shorrocks, A. F. (1982). On the distance between income distributions, *Econometrica* **50**(5): pp. 1337–1339.

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

Van de gaer, D. (1993). *Equality of Opportunity and Investment in Human Capital*, PhD thesis, Katholieke Universiteit Leuven.

Weymark, J. A. (2003). Generalized gini indices of equality of opportunity, *Journal of Economic Inequality* **1**: 5–24. 10.1023/A:1023923807503.

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

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

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

A Definitions and proofs

A.1 Notions of stochastic dominance

Consider the following transformation of the inverse cumulative distribution function $F^{-1}(p)$, indexed by the natural number k:

$$\begin{aligned}\Lambda^k(p) &:= \int_0^p \Lambda^{k-1}(t)\,dt \\ \Lambda^2(p) &:= \int_0^p F^{-1}(t)\,dt\end{aligned}$$

where $\Lambda^2(p)$ is the *generalized Lorenz curve* introduced by Gastwirth (1971). The distribution $F(y)$ *Inverse Stochastic Dominates* $F'(y)$ at order k, denoted $F(y) \succcurlyeq_{ISDk} F'(y)$, if and only if $\Lambda^k(p) \geq \Lambda'^k(p)$ for all $p \in [0,1]$, where the inequality holds strict at least for some values of p. At order $k = 1$, ISD1 and stochastic dominance coincide. In both cases, the dominant distribution can be obtained from the dominated only through increments of incomes. At order $k = 2$, ISD2 and generalized Lorenz dominance coincide. In both cases, the dominant distribution can be obtained from the dominated trough increments and progressive transfers on incomes. For $k \geq 3$, inverse and direct stochastic dominance orders differ. The ISDk induces a partial ranking of distributions: when the curves representing $\Lambda^k(p)$ and $\Lambda'^k(p)$ cross in at least one point, the two distributions F and F' cannot be ordered.

Following Muliere and Scarsini (1989) and Aaberge et al. (2014) it can be established that ISDk is a necessary and sufficient condition for agreement among all preferences in $\mathcal{R}^k$ over the preferred distribution. Furthermore, Maccheroni et al. (2005) show that if $F_\pi(y|c,e) \succcurlyeq_{ISDk} F_\pi(y|c',e)$ then $F_\pi(y|c,e) \succcurlyeq_{ISDl} F_\pi(y|c',e)$, for all $l > k$.[33]

We now show that there always exists a degree k at which any pair of distributions can be compared according to ISD. The dominating distribution is the one that grants higher incomes to the poorest quantiles.

Proposition 7 *For any pair of distributions with bounded support, with inverse cumulative distribution functions denoted by* $F^{-1}(.)$ *and* $F'^{-1}(.)$*, if* $\exists p_\beta > 0 | \forall p \in [0, p_\beta)\ F^{-1}(p) \geq F'^{-1}(p)$ *and the strict inequality holds on a positive mass interval* $[p_\beta - \epsilon, p_\beta)$ *with* $\epsilon > 0$*, we have:*

$\exists k^* \in \mathbb{R}_+$ *and finite such that* $F \succ_{ISDk} F'\ \forall k \in \mathbb{N}_+$ *such that* $k > k^*$*.*

Proof. Using the gap curve notation, we define $\Gamma^k(p) := \Gamma^k(F, F', p) = \Lambda^k(p) - \Lambda'^k(p)$ and $\Gamma(p) := F^{-1}(p) - F'^{-1}(p)$ at any $p \in [0,1]$. Integrating by parts the function $\Gamma^k(p)$,

[33] It is well known (e.g. Muliere and Scarsini 1989) that first and second order inverse stochastic dominance are equivalent to direct first and second order stochastic dominance, which is implemented by generalized Lorenz dominance when means coincide (Shorrocks 1983).Atkinson (1970) showed the logical relation between GL dominance with fixed means and an the utilitarian social welfare function, later generalized to all S-concave social welfare functions and to income distributions with different means.

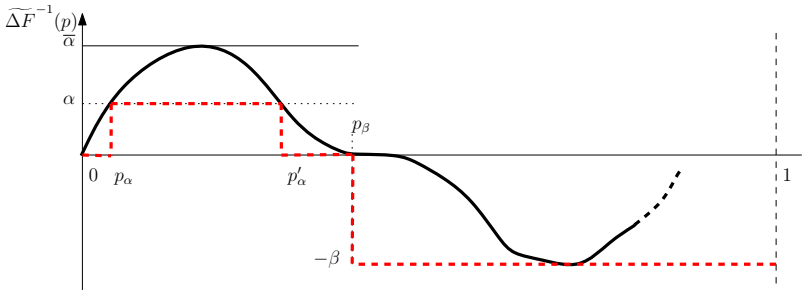


Figure 4: Proof of Proposition 7. The curves $\Gamma(p)$ (solid black) and $\widetilde{\Gamma}(p)$ (dashed red)

up to $k-2$ times, gives:

$$\begin{aligned}
\Gamma^k(p) &= \int_0^p \Gamma^{k-1}(t)dt = -\int_0^p t\cdot\Gamma^{k-2}(t)dt + \Big[t\Gamma^{k-1}(t)\Big]_0^p \\
&= \int_0^p (p-t)\Gamma^{k-2}(t)dt \\
&= \int_0^p \frac{1}{2}(p-t)^2\Gamma^{k-3}(t)dt + \left[\frac{1}{2}(p-t)^2\Gamma^{k-2}(t)\right]_0^p \\
&= \int_0^p \frac{1}{(k-2)!}(p-t)^{k-2}\Gamma(t)dt \qquad (6)
\end{aligned}$$

The result in (6) follows from the fact that $\Lambda^k(0)=0$ and therefore $\Gamma^k(0)=0$ for any k, and that $\Gamma^2(p)=\int_0^p\Gamma(t)dt$.

Let assume from the outset that F should be the dominating distribution. If it is the case, there must exist a point $p_\beta\in[0,1]$ such that $\Gamma(p)\geq 0$ for all $p\in[0,p_\beta]$ with possibly some strong inequalities. Moreover, in the worst case $\Gamma(p)\leq 0$ for all $p\in(p_\beta,1]$. If $\Gamma(p)=0$ for all $p\in[0,p_\beta]$, then by definition $F'\succcurlyeq_{ISD1}F$. Furthermore, since F and F' are left continuous and bounded, then $\Gamma(p)$ is well behaved and takes only finite values on $p=1$ and $p=0$. Though not necessary, this assumption simplifies some technicalities.

This discussion clarifies that the upper bound (denoted $\overline{\alpha}>0$) and lower bound (denoted $-\beta>0$) of $\Gamma(p)$ should exists, as illustrated in figure 4. Formally: $\overline{\alpha}:=\sup\{\Gamma(p):p\in[0,p_\beta)\}>0$ and $-\beta:=\inf\{\Gamma(p):p\in(p_\beta,1]\}<0$.

Consider a scalar $\alpha\in(0,\overline{\alpha}]$ defining *at least* two points $p_\alpha,p'_\alpha\in[0,p_\beta)$, such that $\Gamma(p)>0$ for all $p\in[p_\alpha,p'_\alpha]$. This allows to define a new curve, denoted $\widetilde{\Gamma}(p)$ and represented by a dashed red line in figure 4. This curve equation is:

$$\widetilde{\Gamma}(p):=\begin{cases}0 & \text{if } p\in[0,p_\alpha)\\ \alpha & \text{if } p\in[p_\alpha,p'_\alpha]\\ 0 & \text{if } p\in(p'_\alpha,p_\beta)\\ -\beta & \text{if } p\in[p_\beta,1]\end{cases}$$

The points α and $-\beta$ are identified by a given gap curve $\Gamma(p)$, so it always holds by construction that $\widetilde{\Gamma}(p)\leq\Gamma(p)$ for all $p\in[0,1]$. This is because $\widetilde{\Gamma}(p)$ reduces to the minimum the positive part of $\Gamma(p)$ while it magnifies the part where the curve takes on negative values. As a consequence, any recursive integral of $\widetilde{\Gamma}(p)$ will always lie above the corresponding integral of $\Gamma(p)$. To prove the proposition, it is therefore sufficient to show that, given the premises, there exists an order of integration of $\widetilde{\Gamma}(p)$ for which the resulting

curve is always non-negative for any $p \in [0,1]$. Denote a real positive $\widetilde{k}^*$ such that $\forall k > \widetilde{k}^*$:

$$\int_0^p \frac{1}{(k-2)!}(p-t)^{k-2}\widetilde{\Gamma}(t)dt \geq 0 \quad \forall p \in [0,1], \tag{7}$$

then, there must exists also a real positive value k^* for which the proposition is true.

Noticing that the expression in (7) can take on non-zero values only on the interval $[p_\alpha, p'_\alpha]$ and negative on the interval $[p_\beta, 1]$, it can be simplified as follows:

$$\begin{aligned}\int_0^p \frac{(p-t)^{k-2}}{(k-2)!}\widetilde{\Gamma}(t)dt &= \frac{1}{(k-2)!}\left[\int_{p_\alpha}^{p'_\alpha}(p-t)^{k-2}\alpha dt + \int_{p_\beta}^{p}(p-t)^{k-2}(-\beta)dt\right] \\ &= \frac{\left\{\alpha\left[(p-p_\alpha)^{k-1} - (p-p'_\alpha)^{k-1}\right] - \beta(p-p_\beta)^{k-1}\right\}}{(k-2)!} \geq 0.\end{aligned}$$

To check the validity of the proposition, it is necessary to verify that the above expression is true for any $p \geq p_\beta$ only. To show this, it is sufficient to show that there exists a $\widetilde{k}^*$ such that:

$$\frac{(p-p_\alpha)^{k-1} - (p-p'_\alpha)^{k-1}}{(p-p_\beta)^{k-1}} \geq \frac{\beta}{\alpha}, \quad \forall p \geq p_\beta. \tag{8}$$

By construction of $\widetilde{\Gamma}(p)$, if the condition holds for $p = 1$, then it must hold for all $p < 1$, because the differential takes only negative values for $p \geq p_\beta$. Note that the numerator and denominator of the left hand side of (8) are positive, but the ratio is not said to be greater than one. Nevertheless, one can always pick up a value of $\alpha < \overline{\alpha}$ such that $(p - p'_\alpha) \approx (p - p_\beta)$ and (8) is therefore satisfied if and only if the following holds:

$$\left(\frac{1-p_\alpha}{1-p_\beta}\right)^{k-1} \geq 1 + \frac{\beta}{\alpha}. \tag{9}$$

Both sides of (9) are positives and greater than one. Thus, by taking logs on the left and right side, it is easy to show that the integral condition in (6) is satisfied if and only if the integration order $\widetilde{k}^*$ is large enough to verify:

$$\widetilde{k}^* \geq 1 + \frac{\ln(1+\beta/\alpha)}{\ln(1-p_\alpha) - \ln(1-p_\beta)}.$$

Note that $\widetilde{k}^*$ is positive and greater than one and it always exists *finite* for any $0 < p_\alpha < p_\beta < 1$ and for $\alpha, \beta > 0$. Therefore the value k^* exists as well, which concludes the proof. ■

A.2 Proof of Proposition 1

Proof. By contradiction.

Assume $\exists \tilde{p} \in]0,1[$ such that $|\Gamma(F_1, F'_1, \tilde{p})| > |\Gamma(F_0, F'_0, \tilde{p})|$. For $\pi = 0, 1$, $\Gamma(F_\pi, F'_\pi, p)$ is left continuous since F_π and F'_π are left continuous. Hence, $\exists \epsilon > 0$ such that $\forall p \in [\tilde{p} - 2\epsilon, \tilde{p}]$, $|\Gamma(F_1, F'_1, p)| > |\Gamma(F_0, F'_0, p)|$ and $\text{sign}(\Gamma(F_\pi, F'_\pi, p)) = \text{sign}(\Gamma(F_\pi, F'_\pi, \tilde{p}))$.

Consider the individual preferences $\tilde{W}$ given by the triangular weighting scheme over the interval $[\tilde{p}-2\epsilon, \tilde{p}]$: $\forall p \in [0,1]$, $\tilde{w}(p) = [(\epsilon - |p - (\tilde{p} - \epsilon)|)/\epsilon^2] \cdot \mathbb{1}_{p \in [\tilde{p}-2\epsilon, \tilde{p}]}$, where $\mathbb{1}$ denotes

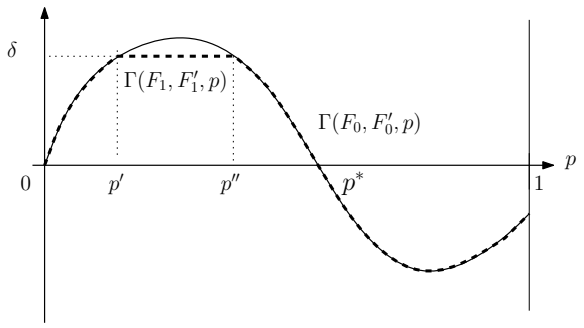


Figure 5: The curves $\Gamma(F_0, F_0', p)$ (solid line) and the perturbation generating $\Gamma(F_1, F_1', p)$ (dashed line).

the indicator function. For preferences $\tilde{W}$, the economic distance in social state π is given by: $|\Delta_{\tilde{W}}(F, F')| = \int_{\tilde{p}-2\epsilon}^{\tilde{p}} \tilde{w}(p)|\Gamma(F, F', p)|dp$. Henceforth $|\Delta_{\tilde{W}}|(F_1, F_1')| > |\Delta_{\tilde{W}}(F_0, F_0')|$ which violates equalization. ■

The reciprocal is not true. Figure 5 provides a counter-example. The plain line gives the gap curve under $\pi = 0$. At value $\bar{p}$, the curve crosses the horizontal axis. Hence, under $\pi = 0$, type c receives higher outcomes than type c' in the bottom of the distribution but lower outcomes in the top. Define the areas $A = \int_0^{p^*} \Gamma(F_0, F_0', p)dp$ and $B = -\int_{p^*}^1 \Gamma(F_0, F_0', p)dp > 0$. Now consider the weighting scheme that gives weight $\alpha \geq 0$ for percentiles below p^* and $\beta \geq 0$ above.[34] The economic distance for this weighting scheme is $|\alpha A - \beta B|$. For α close enough to zero, type c' is preferred to type c and the distance is given by $\beta B - \alpha A$. Under $\pi = 1$, the gap curve is given by the dashed line, which is similar to the plain line except that the advantage of type c has been reduced by a small cumulative amount ϵ in the bottom part of the distribution, so that $\int_0^{p^*} \Gamma(F_1, F_1', p)dp = A - \epsilon < A$. Gap curve dominance is obviously satisfied.[35] At the same time, individuals who initially preferred the distribution of type c' to that of type c will agree that the economic distance between type c and c' has increased and EZOP is thus violated.

A.3 Proof of Proposition 2

Proof. For the *sufficiency* part, assume that $F_0 \succcurlyeq_{ISD1} F_0'$ and $F_1 \succcurlyeq_{ISD1} F_1'$. As a consequence:

$$\forall W \in \mathcal{R}, \forall \pi \int_0^1 w(p)F_\pi^{-1}(p)dp > \int_0^1 w(p)F_\pi'^{-1}(p)dp.$$

Consequently, for all $W \in \mathcal{R}$, we can write:

$$\Delta_W(F_\pi, F_\pi') = \int_0^1 w(p)\Gamma(F_\pi, F_\pi', p)dp.$$

Hence, we have:

$$\Delta_W(F_0, F_0') - \Delta_W(F_1, F_1') = \int_0^1 w(p)[\Gamma(F_0, F_0', p) - \Gamma(F_1, F_1', p)]dp. \tag{10}$$

[34] This weighting scheme is given by $w(p) = \alpha + (\beta - \alpha).\mathbb{1}_{p>p^*}$, with $\alpha, \beta \geq 0$ and $\alpha p^* + \beta(1 - p^*) = 1$.

[35] The same result would obtain even in cases where $A < B$ and the weighting scheme is coherent with risk aversion, i.e. when $\alpha > \beta$.

If $[\Gamma(F_0, F_0', p) - \Gamma(F_1, F_1', p)] \geq 0$ for all p, since the weights $w(p)$ are non-negative, the integrand in equation (10) is positive for all p and the integral is positive.

For the *necessity* part, note that if $[\Gamma(F_0, F_0', p) - \Gamma(F_1, F_1', p)]$ is negative in the neighborhood of a quantile p_0, we can find a weight profile $w(p)$ that is arbitrarily small outside this neighborhood and it makes the integral negative. ■

A.4 Proof of Proposition 3

Proof. We develop the proof for gap curves dominance on the results in Aaberge et al. (2014). As a consequence of the dominance hypothesis, we have:

$$\forall W \in \mathcal{R}^2, \forall \pi \int_0^1 w(p) F_\pi^{-1}(p) dp > \int_0^1 w(p) F_\pi'^{-1}(p) dp.$$

Consequently, for all $W \in \mathcal{R}^2$, we can write:

$$\Delta_W(F_\pi, F_\pi') = \int_0^1 w(p) \Gamma(F_\pi, F_\pi', p) dp.$$

Hence, $\forall W \in \mathcal{R}^2$ we have:

$$\Delta_W(F_0, F_0') - \Delta_W(F_1, F_1') = \int_0^1 w(p)[\Gamma(F_0, F_0', p) - \Gamma(F_1, F_1', p)] dp. \tag{11}$$

It is possible to integrate (11) by parts once,

$$\begin{aligned} \Delta_W(F_0, F_0') - \Delta_W(F_1, F_1') &= w(1) \int_0^1 \left[\Gamma(F_0, F_0', p) - \Gamma(F_1, F_1', p)\right] \\ &+ \int_0^1 (-1) w'(p) \int_0^p \left[\Gamma(F_0, F_0', t) - \Gamma(F_1, F_1', t)\right] dt dp \end{aligned}$$

By $W \in \mathcal{R}^2$ then $w(1) = 0$ and the first term disappears. By $w'(p) \leq 0$ for all p makes $\int_0^p [\Gamma(F_0, F_0', t) - \Gamma(F_1, F_1', t)] \, dt \geq 0 \; \forall p \in [0, 1]$ sufficient for $\Delta_W(F_0, F_0') - \Delta_W(F_1, F_1') \geq 0$ in (11), which gives the EZOP condition. Moreover, Lemma 1 in Aaberge (2009) gives the necessary part. ■

A.5 Proof of Proposition 4

Proof. We develop the proof for gap curves dominance on the results in Aaberge et al. (2014). As a consequence of the dominance hypothesis, we have:

$$\forall W \in \mathcal{R}^k, \forall \pi \int_0^1 w(p) F_\pi^{-1}(p) dp > \int_0^1 w(p) F_\pi'^{-1}(p) dp.$$

Consequently, for all $W \in \mathcal{R}^k$, we can write:

$$\Delta_W(F_\pi, F_\pi') = \int_0^1 w(p) \Gamma(F_\pi, F_\pi', p) dp.$$

Hence, $\forall W \in \mathcal{R}^k$ we have:

$$\Delta_W(F_0, F_0') - \Delta_W(F_1, F_1') = \int_0^1 w(p)[\Gamma(F_0, F_0', p) - \Gamma(F_1, F_1', p)]dp. \tag{12}$$

It is possible to integrate (12) by parts k times,

$$\begin{aligned}\Delta_W(F_0, F_0') - \Delta_W(F_1, F_1') &= w(1)\int_0^1 \left[\Gamma(F_0, F_0', p) - \Gamma(F_1, F_1', p)\right] \\ &+ \sum_{j=1}^{i}(-1)^j \frac{d^j w(1)}{dp^j}\left[\Gamma^k(F_0, F_0', 1) - \Gamma^k(F_1, F_1', 1)\right] \\ &+ (-1)^i \int_0^1 \frac{d^i w(p)}{dp^i}\left[\Gamma^k(F_0, F_0', p) - \Gamma^k(F_1, F_1', p)\right] dp.\end{aligned}$$

By $W \in \mathcal{R}^k$ then $w(1) = 0$ and $\frac{d^j w(1)}{dp^j} = 0$ for all $j \leq i$ and the first and second terms above disappear. The fact that the sign of $(-1)^i \frac{d^i w(p)}{dp^i}$ is always positive by construction makes the conditions $\left[\Gamma^k(F_0, F_0', p) - \Gamma^k(F_1, F_1', p)\right] \geq 0 \ \forall p \in [0,1]$ sufficient for $\Delta_W(F_0, F_0') - \Delta_W(F_1, F_1') \geq 0$ as in (12), which defines the EZOP condition. Moreover, Lemma 1 in Aaberge (2009) gives the necessary part. ■

B Statistical inference for gap curve dominance

B.1 Setting and null hypothesis

Consider a sample $y_1, y_2, \ldots, y_n$ of n draws from a random variable Y with distribution F. Let assume for simplicity that $y_1 \leq y_2 \leq \ldots \leq y_n$, so that y_i refers to the observation in position i in the ranking. The empirical distribution for the sample is denoted $\widehat{F}(y) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(y_i \leq y)$ while the empirical quantile function is denoted $\widehat{F}^{-1}(p) = \inf\{y : \widehat{F}(y) \geq p\}$. If $\widehat{F}$ is a consistent estimator for F, then $\widehat{\Lambda}^k$ is a consistent estimator for Λ^k.

The empirical counterparts of the distributions F_1 and F_1', corresponding to circumstances c and c', are denoted $\widehat{F}_1$ and $\widehat{F'}_1$ respectively, where in general $n_{c,1} \neq n_{c',1}$. These distributions can be obtained from random samples drawn from a well defined population. In the policy evaluation case, in particular, the actual distributions are estimated from the sample of treated units. In the application, we look more specifically at the sample of children born in post reform cohorts in treated municipalities to infer these distributions. In many cases, also $\widehat{F}_0$ and $\widehat{F'}_0$ can be obtained from well defined random samples. In these situations, standard ISDk inference procedures can be adopted to determine the dominating distribution.[36] In the policy evaluation setting, however, $\pi = 0$ represents a counterfactual setting where sample data are not available. The two distributions $\widehat{F}_0$ and $\widehat{F'}_0$ can be estimated by applying to $\widehat{F}_1^{-1}(p)$ and to $\widehat{F'}_1^{-1}(p)$ the corresponding conditional QTE estimates $\hat{QTE}(p|c)$ and $\hat{QTE}(p|c')$ at every $p \in [0,1]$. In this case, inference procedures can be based on bootstrapped covariances and standard errors.

[36] See Beach and Davidson (1983) and Zheng (2002) for estimators of quantile functions and generalized Lorenz functions coordinates, while see Andreoli (2013) for extensions of inference on ISD at orders higher than the second. See also Davidson and Duclos (2000) for inference procedures of direct stochastic dominance analysis.

Whenever $F_\pi \succ_{ISDk} F_\pi$ for all π, the gap curves differences are well defined and gap curves dominance and equality null hypothesis can be stated by setting conditions on $\Lambda^k_\pi(p)$ and $\Lambda'^k_\pi(p)$ in every state π.

$$\begin{array}{ll} H^k_0: & \Lambda^k_0(p) - \Lambda'^k_0(p) \geq \Lambda^k_1(p) - \Lambda'^k_1(p) \quad \text{for all } p \in [0,1]; \\ H^k_1: & \Lambda^k_0(p) - \Lambda'^k_0(p) < \Lambda^k_1(p) - \Lambda'^k_1(p) \quad \text{for some } p \in [0,1]. \end{array}$$

The random process $\Lambda^k(p)$ is, in general, continuous. However, it is more convenient to express equality and dominance null hypothesis as linear equality and inequality constraints on a finite number m of abscissae $p \in \{p_1, \ldots, p_m\}$. The estimates of $\widehat{\Lambda}^k(p)$ corresponding to these abscissae are organized in a column vector of coordinates:

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

with $\mathbf{\Lambda}^k$ being the corresponding vector in the population. Within the discrete setting it can be shown that:

$$\widehat{\mathbf{\Lambda}}^k \text{ is asymptotically distributed as } \mathcal{N}\left(\mathbf{\Lambda}^k, \frac{\mathbf{\Sigma}^k}{n}\right), \tag{13}$$

where we use $\frac{\widehat{\mathbf{\Sigma}}^k}{n}$ as the estimator of the asymptotic $m \times m$ covariance matrix of $\widehat{\mathbf{\Lambda}}^k$. As a consequence of asymptotic normality, test statistics for ISDk and gap curve dominance relations have well known distributional properties.

Asymptotic normality results from applying virtually all estimators proposed in the literature. We retain asymptotic normality also for cases in which the covariances are bootstrapped. A bootstrap estimator consists in randomly sampling from the original sample a sufficiently large amount of time, and to obtain from each bootstrapped sample a set of estimates of the desired coefficients that are stored aside. The bootstrapped coefficients and covariances are the empirical average and covariances obtained from the set of bootstrapped coefficients.

B.2 Application to Gap curves dominance

We estimate dominance conditions for discrete processes, summarized by vectors of coordinates $\mathbf{\Lambda}^k_\pi$ and $\mathbf{\Lambda}'^k_\pi$, corresponding respectively to the population distributions F_π and F'_π in both $\pi = 0$ and $\pi = 1$. We define $\mathbf{\Lambda}^k_\Gamma$ the $4m \times 1$ vector obtained by staking the vectors $\mathbf{\Lambda}^k_0$, $\mathbf{\Lambda}'^k_0$, $\mathbf{\Lambda}^k_1$ and $\mathbf{\Lambda}'^k_1$ in this precise order. The corresponding estimates are collected in the $4m \times 1$ vector $\widehat{\mathbf{\Lambda}}^k_\Gamma$, and we use $n = n_{c,0} + n_{c',0} + n_{c,1} + n_{c',1}$ to denote the overall sample size, gathering together all observations in the sub-samples delimited by circumstances c and c' under $\pi = 0$ and $\pi = 1$, while $r_{c,\pi} = n_{c,\pi}/n$ is the relative size of each sub-sample. In the case of bootstrapped estimators, as in the policy evaluation application, it is sufficient to set $n_{c,\pi} = 1$.

The hypothesis of gap curve dominance can be reformulated as a sequence of m linear constraints on the vector $\mathbf{\Lambda}^k_\Gamma$. Let $\mathbf{R}_\Gamma = (\mathbf{R}, -\mathbf{R})$ be the $m \times 4m$ *difference-in-differences* matrix, where $\mathbf{R} = (\mathbf{I}, -\mathbf{I})$ and $\mathbf{I}$ is an identity matrix of size m. Define the parametric vector $\boldsymbol{\gamma}_k = \mathbf{R}_\Gamma \mathbf{\Lambda}^k_\Gamma$.

We make two (non-testable) assumptions: (i) F_π and F'_π are independent processes for all π; (ii) the independence extends also across policy regimes. This latter assumption

is verified when the sampling scheme is based upon randomized assignment to treatment and control groups. Under the two assumptions of independence and using the result in (13), it holds that:

$$\sqrt{n}\,\widehat{\boldsymbol{\gamma}}_k \;=\; \sqrt{n}\,\mathbf{R}_\Gamma\,\widehat{\boldsymbol{\Lambda}}^k_\Gamma \;\text{ is asymptotically distributed as }\; \mathcal{N}\left(\sqrt{n}\,\mathbf{R}_\Gamma\,\boldsymbol{\Lambda}^k_\Gamma,\; \boldsymbol{\Phi}\right), \qquad (14)$$

where $\widehat{\boldsymbol{\gamma}}_k$ denotes the sample estimate of $\boldsymbol{\gamma}_k$, and

$$\boldsymbol{\Phi} \;=\; \mathbf{R}_\Gamma\,\mathrm{diag}\left(\frac{\boldsymbol{\Sigma}^k_{c,0}}{r_{c,0}}, \frac{\boldsymbol{\Sigma}^k_{c',0}}{r_{c',0}}, \frac{\boldsymbol{\Sigma}^k_{c,1}}{r_{c,1}}, \frac{\boldsymbol{\Sigma}^k_{c',1}}{r_{c',1}}\right)\mathbf{R}^t_\Gamma.$$

The empirical estimator of the asymptotic variance, $\widehat{\boldsymbol{\Phi}}$, is obtained by plugging $\widehat{\boldsymbol{\Sigma}}^k_{c,\pi}$ in the previous formula. We now discuss, within this framework, the test statistics associated to equality and dominance null hypothesis for pairs of gap cures. We conclude the section by showing that the procedure can be simplified by using conditional QTE bootstrapped estimates, as proposed in the empirical session.

B.3 Testing equality in gap curves

The null and alternative hypothesis for equality in gap curves coordinates associated to the set of abscissae $\{p_1, \ldots, p_m\}$ are:

$$H^k_0 :\; \boldsymbol{\gamma}_k = \mathbf{0} \qquad H^k_1 :\; \boldsymbol{\gamma}_k \neq \mathbf{0}.$$

Under the null hypothesis, it is possible to resort to a Wald test static T^k_1:

$$T^k_1 \;=\; n\,\widehat{\boldsymbol{\gamma}}^t_k\,\widehat{\boldsymbol{\Phi}}^{-1}\,\widehat{\boldsymbol{\gamma}}_k.$$

Given the convergence results in (14), the asymptotic distribution of the test T^k_1 is χ^2_m. The p-value tabulation follows the usual rules.

B.4 Testing dominance in gap curves

The null and alternative hypothesis for dominance in gap curves can be reformulated as a sequence of positivity constraints on the vector $\boldsymbol{\gamma}_k$:

$$H^k_0 :\; \boldsymbol{\gamma}_k \in \mathbb{R}^m_+ \qquad H^k_1 :\; \boldsymbol{\gamma}_k \notin \mathbb{R}^m_+$$

The Wald test statistics with inequality constraints has been developed by Kodde and Palm (1986). For this set of hypothesis, the test statistics T^k_2 is defined as:

$$T^k_2 = \min_{\boldsymbol{\gamma}_k \in \mathbb{R}^m_+} \left\{ n\,(\widehat{\boldsymbol{\gamma}}_k - \boldsymbol{\gamma}_k)^t\,\widehat{\boldsymbol{\Phi}}^{-1}\,(\widehat{\boldsymbol{\gamma}}_k - \boldsymbol{\gamma}_k) \right\}.$$

Kodde and Palm (1986) have shown that the statistic T^k_2 is asymptotically distributed as a mixture of χ^2 distributions, provided that (14) holds:

$$T^k_2 \;\sim\; \overline{\chi}^2 \;=\; \sum_{j=0}^{m} w\left(m, m-j, \widehat{\boldsymbol{\Phi}}\right) \Pr\left(\chi^2_j \geq c\right),$$

with $w\left(m, m-j, \widehat{\mathbf{\Phi}}\right)$ the probability that $m-j$ elements of $\boldsymbol{\gamma}_k$ are strictly positive.[37]

To test the reverse dominance order, that is $\Gamma(\Lambda_1^k,\, \Lambda_1'^k, p) \geq \Gamma(\Lambda_0^k,\, \Lambda_0'^k, p)$ for all $p \in [0,1]$, it is sufficient to replace $-\widehat{\boldsymbol{\gamma}}_k$ and $-\boldsymbol{\gamma}_k$ for their positive counterparts.

B.5 Testing equality and dominance using QTE

We propose to assess gap curves dominance for any pair of types $c \neq c'$ by comparing the full distribution of QTE associated to conditional distributions of these two types. When $F_\pi \succcurlyeq_{ISD1} F'_\pi$ for $\pi = 0, 1$, the gap curves dominance condition rewrites:

$$\Gamma(F_0, F'_0, p) - \Gamma(F_1, F'_1, p) \;=\; QTE(p|c') - QTE(p|c) \quad \forall p \in [0,1].$$

Gap curves dominance is equivalently assessed by checking if the conditional QTE distribution of the type which is disadvantaged in every policy regime is larger than the corresponding conditional QTE distribution of the advantaged type. This means that the advantaged type gains less from the policy compared to the disadvantaged type. Since this hold at every conditional quantile of the two types distributions, the disadvantaged type catches up the advantaged one and opportunities are equalized for this pair. The baseline econometric models provides estimates of the QTE at different quantiles of the conditional outcomes distributions, along with their covariances. In the application, these estimators are obtained through bootstrap methods.

Let denote the vector of conditional QTE on the conditional distribution of type c and calculated for a finite number of intercepts $p \in \{p_1, \ldots, p_m\}$ as $\hat{\boldsymbol{\beta}}_c = \left(\hat{QTE}(p_1|c), \ldots, \hat{QTE}(p_m|c)\right)$. The covariance matrix of this estimator is denoted $\boldsymbol{\Sigma}_{QTE|c}$ and is estimated by the empirical covariance of sequence of coefficient estimates bootstrapped on resampled data. In analogy with the previous notation, let $\widehat{\boldsymbol{\beta}}$ be the vector obtained by staking the QTE distributions $\widehat{\boldsymbol{\beta}}_c$ and $\widehat{\boldsymbol{\beta}}_{c'}$ associated to the pair of circumstances c, c'. The m-variate vector of parameters of interest is denoted $\widehat{\boldsymbol{\gamma}}_{QTE} = \mathbf{R}\,\widehat{\boldsymbol{\beta}}$. In analogy with (14), the covariance of $\widehat{\boldsymbol{\gamma}}_{QTE}$ is denoted $\mathbf{\Phi}_{QTE}$ and defined as:

$$\mathbf{\Phi}_{QTE} \;=\; n\,\mathbf{R}\mathrm{diag}\left(\boldsymbol{\Sigma}_{QTE|c}, \boldsymbol{\Sigma}_{QTE|c'}\right)\mathbf{R}^t.$$

The underlying assumption is that the QTE estimates associated to different groups are independent.

In this setting, the hypothesis of equality of QTE, as well as the hypothesis that the QTE distribution of the disadvantage type dominates the QTE distribution of the advantage groups, can be tested as shown in the previous sections, using the test statistics T_1 and T_2. This procedure might not be appropriate in cases where F_π and F'_π are not ordered in the same way according to ISD1 under all policy regimes.

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

Supporting online appendix

C Empirical implementation

This section develops an algorithm that allows to test whether the ex ante EZOP condition is satisfied. Assume that individual outcome and circumstances are observed for a representative sample of the population. The following algorithm operationalizes the ex-ante EZOP criterion. The algorithm proposes null hypothesis for conditional outcome distribution and for gap curves, focussing in particular on equality and dominance null hypothesis. The appropriate inference procedure, tests statistics and their distributions that must be used to test these null hypothesis are treated in appendix B. Valid alternatives to equality and dominance are described in Dardanoni and Forcina (1999).

The algorithm defines a procedure for comparing *pairs of distributions* made conditional on circumstances $c_i, c_j \in C$ with $i, j \in \{1, \ldots, T\}$ the set of population types. In the application, types are defined by intervals of parental resources bounded by deciles of the family labor income, with $T = 10$. Hereby, the algorithm develops a procedure for assessing the most demanding non-anonymous ex-ante opportunity equalization criterion. It can be easily adapted to check equalization with anonymity, or with respect to some reference type. We denote with the scalar $\kappa_{ij}(\pi)$ the minimal degree of ISD at which the two distributions can be ranked in a given social state π. The algorithm establishes the steps that must be iterated to determine empirically $\kappa_{ij}(\pi)$. Inference on this scalar remains beyond the scope of this work.

Algorithm 1 (Implementable ex ante EZOP for two types) *The following sequence of estimations and tests implements EZOP:*

(i) $\forall c_i \in C$, $\forall \pi$, *estimate* $F_\pi^{-1}(p|c_i)$ *and its integrals* $\Lambda_\pi^k(p|c_i)$.

(ii) For each (c_i, c_j, π) *with* $i, j \in \{1, \ldots, T\}$ *compute* $\kappa_{ij}(\pi)$ *as follows:*

- *(a) Consider* $k \in \mathbb{N}_{++}$, *with* $k = 1$ *for the first iteration;*
- *(b) Given* k, *define and test the following pair of null hypothesis:*
  $\{H_0 : F_\pi(y|c_i) \succ_{ISDk} F_\pi(y|c_j)$ *vs.* $H_a : F_\pi(y|c_i) \not\succ_{ISDk} F_\pi(y|c_j)\}$
  and
  $\{H_0 : F_\pi(y|c_j) \succ_{ISDk} F_\pi(y|c_i)$ *vs.* $H_a : F_\pi(y|c_j) \not\succ_{ISDk} F_\pi(y|c_i)\}$.
- *(c) Define* $I_k = (a, b)$ *the result of this pair of tests, where* a, b *is equal to 1 if the null hypothesis is rejected and 0 otherwise, respectively for both null hypothesis.*
- *(d) Compute* I_k*:*
  - *if* $I_k = (0, 0)$*:* $\kappa_{ij}(\pi) = \infty$ - **stop**.
  - *if* $I_k = (0, 1)$ *or if* $I_k = (1, 0)$*:* $\kappa_{ij}(\pi) = k$ - **stop**.
  - *if* $I_k = (1, 1)$*: let* $k = k + 1$ *and* **iterate** *from step (b).*

(iii) Define $\kappa_{ij} := \max_\pi\{\kappa_{ij}(0), \kappa_{ij}(1)\}$.

(iv) Verify gap curve dominance at order κ_{ij}, *where* c *and* c' *represent respectively the dominating and dominated distribution out of the pair* c_i, c_j*:*
$\big\{H_0 : \Gamma\big(\Lambda_0^{\kappa_{ij}}(p|c), \Lambda_0^{\kappa_{ij}}(p|c')\big) \geq \Gamma\big(\Lambda_1^{\kappa_{ij}}(p|c), \Lambda_1^{\kappa_{ij}}(p|c)\big)\ \forall p \in [0, 1]$ *and* $H_a :$ H_0*is false*$\big\}$.

- *If* $\kappa = 1$*:*

  - *If H_0 accepted: EZOP is verified.*
  - *If H_0 accepted with equality: neutrality is verified.*
  - *If H_0 rejected: inconclusive, EZOP is rejected.*
- *If $\kappa \geq 2$:*
  - *If H_0 accepted: Necessary conditions for equalization are satisfied.*
  - *If H_0 accepted with equality: neutrality is verified.*
  - *If H_0 rejected: inconclusive, EZOP is rejected.*

When $k = 1$, gap curve dominance can be inferred from comparisons of QTE. Provided that $F_\pi(y|c) \succcurlyeq_{ISD1} F_\pi(y|c')$, then the gap curve dominance relation

$$\Gamma(F_0(y|c), F_0(y|c'), p) - \Gamma(F_1(y|c), F_1(y|c'), p) = F_1^{-1}(p|c') - F_0^{-1}(p|c') - \big(F_1^{-1}(p|c) - F_0^{-1}(p|c)\big)$$

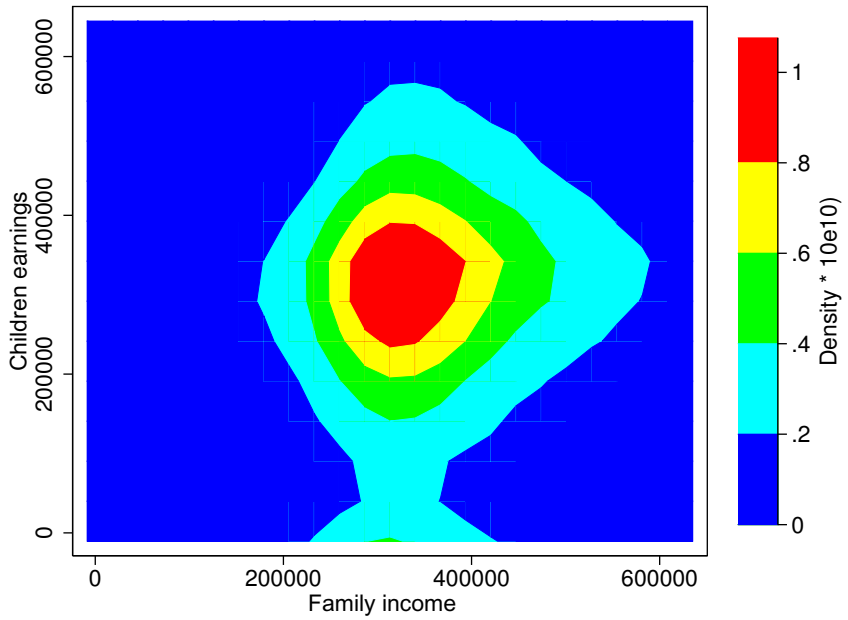
is, by definition of the QTE on conditional types distribution, equivalent to verify that $QTE(p|c') \geq QTE(p|c)$ for every conditional quantile $p \in [0, 1]$. This result allows to use more precise estimators of the gap curves, although the equivalence is valid exclusively in the case $k = 1$.

D Additional material for the empirical application

D.1 Support analysis for QTE estimation

The data used in this paper have been presented in detail by Havnes and Mogstad (2011, 2014). We address the reader to these paper to gather explanations for: (i) the sample composition; (ii) the effects of the Kindergarten Act on child participation in early education and on the effect of the policy on the labor supply of the mothers; (iii) the similarity between treated and control groups in terms of the characteristics of the municipalities they are associated to; (iv) the validity of the DiD identification strategy; (v) the computation of QTE of the policy using RIF-DiD, OLS and change-in-changes estimators; (vi) the definition and distribution of family earnings, along with the effect of the policy on intergenerational earnings elasticity.

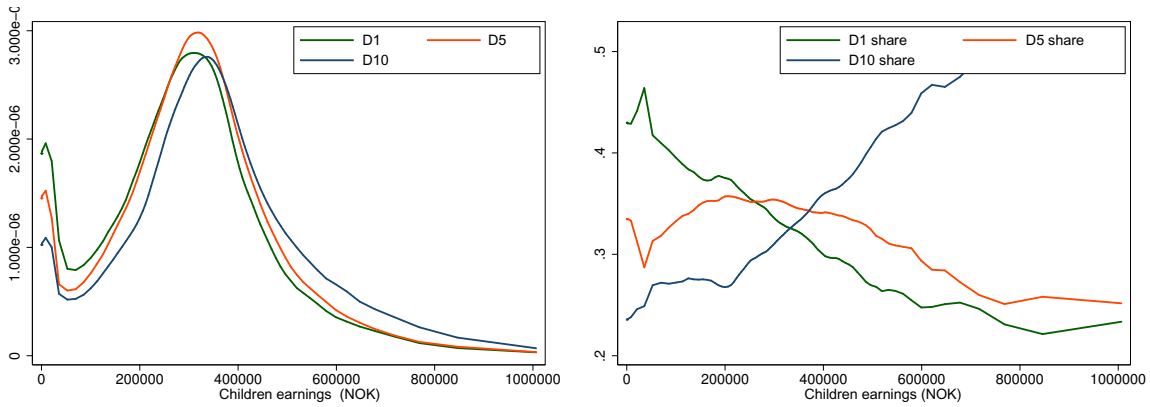
Figure 6 reports a non-parametric estimates of the joint distribution of child earnings and family labor incomes. The high symmetry of the cumulative distribution function ensure that there is enough support for the estimation of QTE for high quantiles of earnings even for the groups coming from less advantaged backgrounds, when running RIF-DiD estimations. The support requirements are satisfied at the bottom of the children earnings distribution, where the density of family labor incomes is sufficiently high for all groups, as confirmed by panel (a) of figure 7 for selected family labor income decile groups. The issues related to lack of support at the top of the children earnings distributions are of marginal relevance. Panel (b) of figure 7 shows in fact that, although the share of group D10 (10th family labor income decile) is proportionally higher than the one of groups D1 or of D5 for top percentiles of the children earnings distributions, still each of these two groups represents half of the group D10 share at that specific percentile. Furthermore, we run conditional QTE models, so the estimates of the parameters $\beta_3 \cdot g(\cdot)$ in (5) for very high income thresholds are mostly associated to top family labor income deciles groups.



t

Figure 6: Joint distribution of child earnings and family labor income.

Note: Distribution is estimated using an epanechnikov kernel with Silverman ROT bandwidth, multiplied by 10e10 for notational simplicity.



(a) Density of child earnings

(b) Population shares by child earnings

Figure 7: Distribution of child earnings by selected family labor income deciles groups.

Note: D1, D5 and D10 refer to children with parental earnings in the first, fifth, and tenth decile, respectively. Density estimates are calculated using an epanechnikov kernel with Silverman ROT bandwidth. Panel (b) refers to shares in the population of children in these three deciles, estimated as the fraction of estimated kernel density to the sum of estimated kernel densities.

D.2 EzOP test: higher order comparisons for selected decile groups

The tests reported in figure 2 show that equalization cannot be rejected for a large majority of the groups involved in the gap curve dominance comparisons. This is because the QTE associated to groups that are relatively advantaged (both in actual and counterfactual settings) are nowhere higher than the QTE of groups that are always relatively disadvantaged.

In *one* out of the 45 cases, involving groups D3 and D5, it is not possible to establish the direction of this dominance relation, since both equalization and disequalization cannot be rejected, while this is not the case for equality of gap curves is rejected. In other *six* cases involving the comparisons of groups D1 versus D4-D9, the joint dominance test rejects that the gap curve in the counterfactual setting lies either above or below the gap curve in the actual setting. In these situations, there are high chances that gaps curves at higher orders are statistically similar and neutrality cannot be rejected.

To complete the equalization of opportunity test, we now apply tests at higher orders of dominance for conditional cdfs and gap curves to a selected number of groups identified by the graphs in figure 2. We test ISD2 relations first, by looking at dominance relations between generalized Lorenz dominance across groups in the actual and counterfactual settings. Results are reported in panels A and B of table 2. For most of the groups considered here, it is possible to conclude with Lorenz curves that group D1 is substantially disadvantaged compared to the others. The generalized Lorenz curves of groups D3 and D5, instead, coincide in the actual setting. In the actual setting, some of the p-values cannot be computed. We rely nevertheless on the fact that the estimated tests statistics are quite large to conclude that some of the null hypothesis are rejected by the data.

Panel C of the figure elaborated equalization tests. The null hypothesis of neutrality of the policy should be accepted at 1% in comparisons involving groups D1 and D4, while both equalization and disequalization between groups D3 and D5 must be accepted, indicating that higher order dominance analysis is needed.

The equalization tests for some of the comparisons involving group D1 are not conclusive, while for many of these comparisons it is not possible to obtain reliable measure of the p-values associated to joint dominance tests. We conclude that within the set of preferences $\mathcal{R}^2$ it is not possible to assess if the gap between groups D1-D5, D1-D6, D1-D7, D1-D8 and D1-D9 has been reduced by the Kindergarten Act. A better tailored inference procedure is needed to tackle cases such as this.

By further restricting consensus on equalization to the class $\mathcal{R}^3$, implemented through dominance tests for Λ^3 curves and their associated gap curve, results are more clearcut. The joint tests reported in table 3 allow to conclude in favor of neutrality of the policy with respect to the gap between groups D3 and D5. Though the analysis of p-values remains not reliable, the Wald joint tests for disequalization of the opportunity profiles between group D1 and the rest, are all zero. Figure 8 motivates that in large part these results are lead by disequalization of opportunity taking place at the bottom of the group D1 children earnings distributions.

Table 2: Equalization of opportunity for selected family labor income deciles groups

	Pairwise groups comparisons: D1 vs.						D3 vs.
	D4	D5	D6	D7	D8	D9	D5
A - Counterfactual generalized Lorenz, dominance relations ($\pi = 0$)							
$H_0 : \sim$	118.9	155.8	280.4	423.9	702.0	851.7	44.4
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
$H_0 : \succcurlyeq$	118.9	155.8	280.4	423.9	702.0	851.7	43.9
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$H_0 : \preccurlyeq$	0.0	0.0	0.0	0.0	0.0	0.0	4.4
	[0.769]	[0.781]	[0.778]	[0.776]	[0.782]	[0.781]	[0.600]
B - Actual generalized Lorenz, dominance relations ($\pi = 1$)							
$H_0 : \sim$	129.1	112.3	213.4	261.3	481.0	494.1	21.7
	[.]	[0.000]	[.]	[.]	[.]	[.]	[0.549]
$H_0 : \succcurlyeq$	129.1	112.3	213.4	261.3	481.0	494.1	21.7
	[.]	[0.000]	[.]	[.]	[.]	[.]	[0.318]
$H_0 : \preccurlyeq$	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	[0.880]	[0.880]	[0.880]	[0.883]	[0.885]	[0.882]	[0.883]
C - Gap curve equality and dominance tests at order ISD2 ($\pi = 0$ vs $\pi = 1$)							
H_0 : Neutrality	43.1	58.2	87.1	110.4	163.0	193.0	60.9
	[0.036]	[0.003]	[0.000]	[0.000]	[.]	[.]	[0.002]
H_0 : Equalization	49.4	59.0	75.4	134.4	104.8	125.7	0.5
	[0.007]	[0.001]	[0.000]	[.]	[0.000]	[.]	[0.858]
H_0 : Disequalization	65.4	99.4	134.4	217.4	226.0	279.8	61.3
	[0.007]	[0.001]	[0.000]	[.]	[0.000]	[.]	[0.857]

Note: generalized Lorenz dominance joint tests on ventiles of children earnings under the counterfactual (panel A) and the actual (panel B) settings, along with gap curve dominance at order 2. Gap cures are defined according to the order of groups defined in panel A and B. Covariances are bootstrapped.

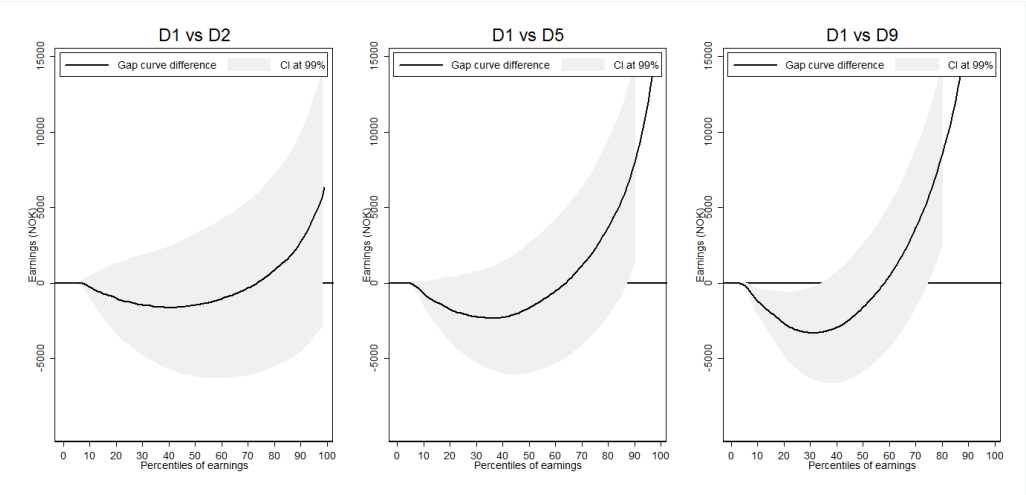


Figure 8: Comparisons of generalized Lorenz curves through gap curves at order 2 for selected pairs of groups

Note: Bootstrapped standard errors.

Table 3: Equalization of opportunity for selected family labor income deciles groups

	Pairwise groups comparisons: D1 vs.						D3 vs.
	D4	D5	D6	D7	D8	D9	D5
A - Counterfactual Λ^3 curves, dominance relations ($\pi = 0$)							
$H_0 : \sim$	111.3	149.9	287.9	407.8	660.5	787.1	34.1
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.018]
$H_0 : \succcurlyeq$	111.3	149.9	287.9	407.8	660.5	787.1	37,016.7
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$H_0 : \preccurlyeq$	0.0	0.0	0.0	0.0	0.0	0.0	37,208.1
	[0.666]	[0.672]	[0.673]	[0.675]	[0.671]	[0.670]	[0.000]
B - Actual Λ^3 curves, dominance relations ($\pi = 1$)							
$H_0 : \sim$	127.1	98.9	225.8	253.2	426.8	479.5	18.4
	[.]	[0.000]	[.]	[.]	[.]	[.]	[0.705]
$H_0 : \succcurlyeq$	127.1	98.9	225.8	253.2	426.8	479.5	18.4
	[.]	[0.000]	[.]	[.]	[.]	[.]	[0.451]
$H_0 : \preccurlyeq$	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	[0.816]	[0.816]	[0.816]	[0.819]	[0.820]	[0.819]	[0.819]
C - Gap curve equality and dominance tests at $k = 3$ ($\pi = 0$ vs $\pi = 1$)							
H_0 : Neutrality	36.3	49.2	74.7	96.7	128.9	159.8	6.5
	[0.098]	[0.013]	[0.000]	[0.000]	[.]	[.]	[0.998]
H_0 : Equalization	36.3	49.2	74.7	96.7	128.9	159.8	0.0
	[0.059]	[0.008]	[0.000]	[0.000]	[.]	[.]	[0.813]
H_0 : Disequalization	0.0	0.0	0.0	0.0	0.0	0.0	6.5
	[0.060]	[0.008]	[0.000]	[0.000]	[.]	[.]	[0.814]

Note: Λ^3 dominance joint tests on ventiles of children earnings under the counterfactual (panel A) and the actual (panel B) settings, along with gap curve dominance at order 3. Gap cures are defined according to the order of groups defined in panel A and B. Covariances are bootstrapped.