Is Equal Opportunity Different from Welfarism?

Gabriel Carroll, University of Toronto
gabriel.carroll@utoronto.ca

February 19, 2025

Abstract

Equal opportunity, widely invoked in popular discourse as a goal for policy, seems at odds with the welfarist approach that is standard in economics. But are they really different? We consider a canonical class of resource allocation problems and ask whether the allocations chosen by an equal-opportunity criterion could also have been chosen under some welfarist criterion. Typically, no such welfarist criterion exists. However, for a rich class of problem specifications, it does exist, and we characterize this class. When the welfarist criterion does exist, it can use either the sum or the min to aggregate individual welfares; the freedom to use more exotic aggregators does not expand the possibilities.

This paper has benefited from discussions and comments from (in random order) Alexis Ghersengorin, Jean Baccelli, and audiences at Warwick, Queen Mary, Bristol, and the 15th Oxford Workshop on Global Priorities Research. The author thanks the economics department at Oxford for their hospitality during a sabbatical.

1 Introduction

How should the conflicting interests of heterogeneous agents be combined to make policy recommendations? In academic economics research, the overwhelming norm is to use welfarist criteria. That is: for each agent in the economy, we posit some measure of the agent's welfare induced by the policy; these individual welfare measures are aggregated together in some way (usually just by summing them together) to form a measure of social welfare; this social welfare measure is then used to evaluate each of the available policy options, and the policy that maximizes social welfare is deemed optimal.

However, many people—both policymakers and laypeople—have normative intuitions about policy choice that are not naturally expressed in welfarist terms (see, for example, the literature discussed in Fleurbaey and Maniquet (2018)). This suggests that economic analysis can better contribute to policy debates if it engages with some of these alternative criteria.

In this paper, we examine one such normative criterion: *equality of opportunity*. This is a concept that is widely invoked in popular discourse and also has inspired a rich literature in political philosophy. There have been several approaches proposed in economics to formalizing this concept (discussed further below). We adopt here a version of the formulation by Roemer (1993, 1998), which has the virtue of being applicable to a wide range of policy questions, and has inspired a significant empirical literature measuring inequality of opportunity. We do not question here whether equality of opportunity (or, indeed, any kind of equality) is actually the appropriate goal for society. Instead, we take as given that the existing widespread interest in equal opportunity provides sufficient motivation for better understanding what it entails.

Given the prevalence of welfarism in economics, arguably the first question for studying equal opportunity (or any other non-welfarist criterion) is: Is it actually distinguishable from welfarism? That is: Consider the policy choices of a planner who optimizes equal opportunity. Could these same choices also have been produced by a welfarist planner? And if not, can we give a simple explanation for what sets them apart? This paper aims to explore this topic.

To make the question concrete, we need to do three things: we need to clarify what is meant by welfarism; we need to explain what is meant by equal opportunity; and we need to delineate a specific class of policy questions to consider. We will now overview each of these in turn.

For purposes of this paper, welfarism means the following. For each agent $i = 1, \dots, n$ in the economy, the effect of any policy (call it ϕ) on agent i's well-being is measured by some quantity $U_i(\phi)$. These welfare levels are then fed into some aggregator function W. Thus, the policy ϕ is then evaluated by $W(U_1(\phi), \dots, U_n(\phi))$, so that the policy that maximizes this quantity is considered best. We assume that W is symmetric in its arguments; thus any heterogeneity among agents (due to preferences, endowments, etc.) is captured by the dependence of $U_i(\phi)$ on i. In practice, W is usually taken to be the sum function (and the welfare criterion is "utilitarian"), but it need not be; somtimes,

W is instead taken to be the min (the "egalitarian" criterion); other aggregators, such as the median, also occur. Much of the academic literature also considers "Pareto-weighted sums" of the form $\sum_i \lambda_i U_i(\phi)$ (where λ_i are individual-specific constants); this might seem to be excluded due to the symmetry requirement, but actually it too is welfarist by this definition, since it can be written as the symmetric aggregate $\sum_i \widetilde{U}_i(\phi)$ where $\widetilde{U}_i = \lambda_i U_i$.

The theory of equal opportunity, in turn, has its philosophical roots in Rawls (1971) and Dworkin (1981a,b). (See the survey by Roemer and Trannoy (2016) for a brief overview of other major contributions in the philosophy literature.) This theory is egalitarian in spirit, but only some inequalities are considered undesirable. Thus, if one person earns less than another because she was born into a poor family, or is a member of a group that faces discrimination, this inequality is considered undesirable, and policy should aim to mitigate it; but if the differences come from different preferences over consumption and leisure, or choosing different kinds of careers in the knowledge that the market rewards them differently, this difference in outcomes is acceptable. More specifically, differences among agents can be described by two kinds of variables, denoted by c and e, and policy should aim to compensate differences in outcomes associated with variation along the c dimension but not the e dimension. To connect explicitly to the concept of opportunity, we can interpret c as capturing variation that leads to unequal opportunity initially, and e as capturing whatever heterogeneity leads to different outcomes given the same opportunity. The question of exactly what real-world attributes correspond to each variable is outside the formal model; we simply take the parameterization of agents by (c, e) as given. (These letters come from the terms "circumstances, effort," which have become standard in the relevant literature. However, it might be better to think of e as "merit" or "skill" rather than "effort," since it is a characteristic of the agent, not a choice variable.)

Any policy ϕ leads to an outcome for each agent, measured numerically by some quantity $A(c, e, \phi)$. This A is called "advantage" or "achievement," and is treated as interpersonally comparable. Note that A need not be identified with individual welfare (this distinction is natural given the philosophical pedigree of the theory, which views "genuine" welfare as interpersonally non-comparable). Instead, A is the object for which we wish to equalize opportunity: for example, we can speak about "equal opportunity for income," or, in a health policy context, "equal opportunity for longevity."

The evaluation of policies is then based on two core principles:

- Aversion to inequality due to differences in c: for agents with different c but the same e, we evaluate policies based on the advantage of the worst-off, $\min_c A(c, e, \phi)$ (the "compensation principle").

- Neutrality toward inequality due to e: for agents with different e but the same c, we evaluate policies based on total advantage, $\sum_e A(c,e,\phi)$ (the "reward principle").

Of course, in general a population will feature differences in both c and e, and no one policy is optimal by both principles simultaneously, so we must mediate between them. The criterion originally proposed by Roemer (1993, 1998) is

$$\sum_e \left(\min_c A(c,e,\phi) \right). \tag{1.1}$$

This criterion can be loosely interpreted as a total measure of the opportunities that are available to everyone in society. Van de gaer (1993) proposed the natural alternative

$$\min_c \left(\sum_e A(c,e,\phi) \right). \tag{1.2}$$

Both of these criteria coincide with the compensation and reward principles in the case of "one-dimensional" populations (all agents have the same c or the same e). These two criteria were originally proposed on the basis of convenience; they, and others, have since been characterized axiomatically by Ooghe, Schokkaert and Van de gaer (2007). For our purposes, we will treat the compensation principle and reward principle—which determine how policies should be chosen in the one-dimensional case—as being central to the theory of equal opportunity. We will not take a firm stand on what criterion is used for general populations, though both (1.1) and (1.2) will be considered.

Thus, equal opportunity resembles welfarism in that it proceeds from a numerical measure for each agent and aggregates them to a population-level objective; but the individual measure need not be interpreted as welfare, and the aggregator does not treat all agents symmetrically but rather distinguishes between aggregation along the c and the e dimensions.

Next, we should identify a specific class of policy questions. This paper will focus on *distribution problems*, in which a divisible good is available, in some specified quantity, to divide among a population of agents. (For a concrete application, we can think of a nonprofit foundation or government agency allocating scholarship funds to university students, with a dual mandate: to use the money productively by giving it to the strongest students, but also to promote social equity by giving extra support to applicants from disadvantaged backgrounds. Thus, c could measure, say, parents' income, while e would be a measure of academic merit.) There are a couple of natural reasons to focus on this class.

First, it figures prominently in much of the economics literature on equal opportunity, such as Roemer (1998) and Maniquet (2004) (discussed further below). Second, one criticism sometimes voiced against non-welfarist criteria is that they might choose Pareto-dominated policies (Kaplow and Shavell, 2001). By studying distribution problems, we can steer clear of that concern, since none of the options is Pareto-dominated.

Let us illustrate the framework with a example taken from Fleurbaey and Maniquet (2011).[1] Suppose that c and e are measured by positive numbers, and suppose the advantage derived by any agent, A, depends on the agent's characteristics (c, e) and the amount of the good x that the agent receives, via the formula[2]

$$A = (x + c)e. \tag{1.3}$$

Consider a population of four agents, where c and e can each take the value 1 or 3, and each combination is represented by one agent.

	$e = 1$	$e = 3$
$c = 1$	undeserving and poor	deserving and poor
$c = 3$	undeserving and rich	deserving and rich

Suppoose there are 16 units of the resource to allocate among the agents. One can calculate that the division that maximizes criterion (1.1), and the resulting levels of advantage A, are given by

x:

	$e = 1$	$e = 3$
$c = 1$	0	9
$c = 3$	0	7

A:

	$e = 1$	$e = 3$
$c = 1$	1	30
$c = 3$	3	30

If the planner instead uses criterion (1.2), the solution is given by

x:

	$e = 1$	$e = 3$
$c = 1$	0	28/3
$c = 3$	0	20/3

A:

	$e = 1$	$e = 3$
$c = 1$	1	31
$c = 3$	3	29

In either case, we see the combination of "compensation" and "reward" features. The

[1] The example also appears quoted in Roemer and Trannoy (2016), although they give the wrong answer for the policy that maximizes criterion (1.1).

[2] Henceforth, we write A as a function of c, e, x, rather than c, e, ϕ as before; this reflects an assumption that an agent's advantage does not depend on the amount of the good received by others.

reward aspect assigns all of the good to the high-e agents, who are most productive at converting it to advantage. The compensation aspect assigns more of the good to low-c than to high-c agents (but the two criteria differ in exactly how much more).

Now we can turn to the topic of this paper. Suppose that we observe the choices made by an equal-opportunity-maximizing (henceforth "opportunitarian") planner—not in this example, but in any population of (c, e) parameters, and with any total amount of the good. In particular, the planner adheres to the compensation principle and the reward principle as above. (In the formal statements, we will restrict ourselves to one-dimensional populations, where these two principles are sufficient to define the planner's preferences.) Are these same choices also consistent with a welfarist criterion?

The space of possible welfarist criteria is large. Many forms for the individual welfare measure U are possible: one could take $U = A = (x + c)e$, but one could also take $U = x + c$, say, or even just $U = x$. That is, different welfarist planners could have different views as to what is the right way to take the individual characteristics (c, e) into account (if at all) in making interpersonal comparisons of welfare; our only restriction is that U should represent each individual's "preferences" by being increasing in x. Planners could also have different views as to what the aggregator function W should be. Rather than make presumptions about these views, our question is whether *any* such view would replicate the opportunitarian choices.

Our first major result, Theorem 3.1, says that, for the advantage function (1.3) as in the example, the answer is no (subject to some mild regularity conditions on the welfarist representation). Moreover, this negative answer holds even if we only consider the opportunitarian's choices on "minimal" populations—consisting of just two agents.

However, this example has a feature that is arguably rather special: a discontinuity in the equal-opportunity solution. For two agents with equal e and just slightly different c's, the good should be distributed approximately equally. Yet, for two agents with equal c and slightly different e's, the linearity of $\sum A$ leads to a corner solution—giving all of the good to the higher-e agent. It is perhaps no surprise, then, that no welfarist criterion can match both kinds of choices.

If we consider instead an advantage function that is even slightly concave in x, this discontinuity disappears: for populations of similar agents, total advantage and worst-off advantage are both maximized by distributing the good roughly equally. For example, suppose that instead of (1.3), advantage is measured by

$$A = \sqrt{(x + c)e}. \tag{1.4}$$

Now there is indeed a welfarist criterion that replicates the opportunitarian's choices on any minimal population (and indeed any one-dimensional population): namely, the total advantage, $\sum A$. Indeed, it is obvious that this criterion satisfies the reward principle. To check the compensation principle, note that if the population has no variation in e, then maximizing total advantage is equivalent to maximizing $\sum \sqrt{x+c}$, which by concavity leads us to equate $x+c$ across all agents (to the extent possible; it may happen that some agents with high c end up at a corner, $x=0$), and this is indeed the same allocation the opportunitarian would choose.

This example shows that, on a qualitative level, the central principles of equal opportunity—the compensation principle and the reward principle—are indeed jointly compatible with a welfarist framework. It also raises the natural question: is the example typical? That is, if we consider an arbitrary specification of the advantage function $A(c,e,x)$, does the situation look more like the example (1.3) or like (1.4)?

A natural conjecture is:

- (Conjecture A) As long as A is strictly concave in x, there exists a welfarist criterion that replicates the opportunitarian's choices.

As it turns out, this conjecture is strongly false. In a suitable sense, for "most" specifications of the advantage function, this replication is impossible. One simple counterexample is given by

$$A = \sqrt{(c+e)x}. \tag{1.5}$$

(The impossibility in this example follows from the general result that will be described shortly.)

On the other hand, the specification (1.4) is not totally unique either; it belongs to a broader family of specifications for which there is a welfarist criterion that replicates the opportunitarian's choices, and we can ask whether the structure of the criterion in the example is typical within that family. In particular, two more natural conjectures are:

- (Conjecture B) In the cases where a welfarist criterion can replicate the opportunitarian's choices, this can be done with A itself being the individual welfare measure.

- (Conjecture C) In the cases where a welfarist criterion can replicate the opportunitarian's choices, this can be done with the criterion being utilitarian (i.e. the aggregator W is the sum function).

As it turns out, Conjecture B is again false. An example is

$$A = 1 - \exp(-cex). \tag{1.6}$$

Here, the opportunitarian's choices cannot be replicated by a welfarist criterion where $U = A$ (we will explain why in Section 4). But they can be replicated by the criterion of maximizing total welfare, where the individual welfare measure is

$$U = \frac{1}{c}(1 - \exp(-cex)) = \frac{A}{c}.$$

They can also be replicated by the criterion of welfare of the worst-off, where now the individual welfare measure is

$$\widetilde{U} = \frac{1}{e}\exp(cex) = \frac{1}{e(1-A)}.$$

Indeed, note that $\partial U/\partial x = 1/\widetilde{U}$, so that (by first-order conditions) the problem of maximizing $\sum U$ is equivalent to equalizing $1/\widetilde{U}$, or equivalently equalizing $\widetilde{U}$. It is then immediate that, in a population with e constant, this is equivalent to equalizing A (the compensation principle). And in a population with c constant, maximizing $\sum U$ is clearly the same as maximizing $\sum A$.

Conjecture C, on the other hand, is true. That is, in looking for a welfarist criterion that replicates the opportunitarian choices, granting the freedom to use more exotic aggregators rather than the sum function actually makes no difference. Moreover, the same is true if we wish to use the min aggregator instead of the sum; as we just saw in the example (1.6), a min-based representation can describe the same choices as the sum-based representation.

All these findings are consequences of our two main results, Theorems 4.4 and 4.5, which identify the class of specifications of A for which a welfarist criterion can replicate the opportunitarian choices. An intuition can be gained by picturing a three-dimensional (c, e, x) space, as shown in Figure 1(a). Holding e fixed, and varying c and x so that the value of A stays constant, traces out a curve in this space (an "A-curve"). The compensation principle says that, in a population with constant e, the good should be allocated to put everyone on the same A-curve (aside from corner solutions). Likewise, holding c fixed, and varying e and x so that the *marginal* advantage $\partial A/\partial x$ stays constant, traces out another curve (a "B-curve"), and the reward principle implies that, in a population with constant c, we should put everyone on the same B-curve. The figure shows a typical point (the hollow circle) and the A-curves and B-curves passing through it.

Suppose the space can be sliced into surfaces ("foliated") in such a way that each surface is a union of A-curves, and also a union of B-curves. (An example of this situation is shown in Figure 1(b); the figures depicts several of the surfaces, and some representative curves on one of the surfaces.) Then, both principles say that all agents should be put on the same surface. This choice can be represented by a welfarist criterion: let U be a function that indexes the surfaces, and evaluate any allocation by the minimum of U in the population. Alternatively, let $\widetilde{U}$ be a function satisfying $\partial\widetilde{U}/\partial x = Z \circ U$, where Z is any decreasing function, and then the criterion of maximizing the sum of $\widetilde{U}$ will also select this allocation. This is the content of Theorem 4.5.

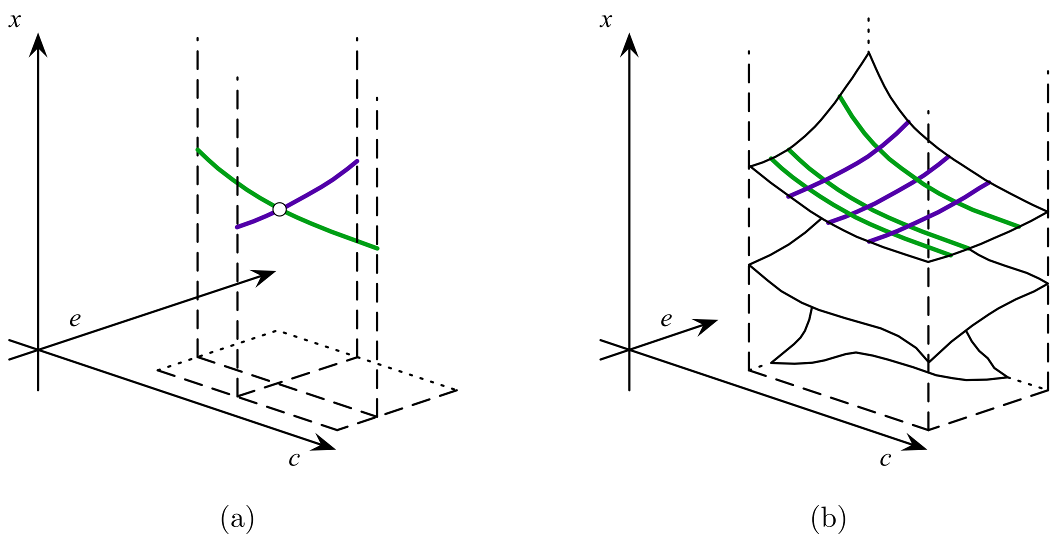

Figure 1: (a) A typical A-curve and B-curve. (b) A foliation in which each surface can be partitioned into A-curves, and also into B-curves.

Conversely, Theorem 4.4 says that existence of such a foliation is also a *necessary* condition for the desired welfarist representation. It is intuitive that this should be a necessary condition if we assume that the welfare aggregator takes either the min or the sum form: either the level sets of U or those of $\partial U/\partial x$, respectively, form the foliation surfaces. What is more surprising is that the foliation remains a necessary condition even without making such an assumption. Roughly speaking, these forms for the aggregator are forced on us by the first-order conditions for welfare maximization.

We can build on these findings to revisit the question of which allocation should be preferred when the population is not one-dimensional—whether we should use criterion (1.1), (1.2), or some other. In the cases where a welfarist criterion can replicate equal opportunity for one-dimensional populations, it also provides a natural choice of allocation for more general populations: namely, putting all agents on the same surface of the foliation. This "canonical" allocation coincides with the solutions chosen by (1.1) and (1.2) when all agents receive positive quantities, but it can differ from them in the corner cases when some agents receive zero. This latter property is not a failing of these specific criteria, however: Proposition 5.7 shows that *no* aggregation procedure that uses only information about the agents' advantage levels and their positions in (c, e) space can select the canonical allocation whenever it exists.

This completes the preview of our results. We defer further interpretive discussion to the conclusion.

1.1 Related literature

The starting point for this work is that the currently prevailing approach to evaluating policies in economics is welfarist. Although this claim might seem uncontentious as a description of the status quo, note that many applied economic questions are not concerned with heterogeneity among agents and so can be studied with models where all agents are identical (at least ex ante), in which case welfarism seems relatively uncontroversial. What is more specifically relevant here is that even in models with heterogeneous agents, it is common practice for the modeler to impose a specific functional form that makes preferences comparable across agents and then to aggregate. A few recent examples of this practice are Ales and Sleet (2022) in a taxation setting, Diamond (2016) in urban economics, and Nigai (2016) in trade.

Perhaps most relevant is the work by Lockwood and Weinzierl (2015), who study an optimal taxation problem in which agents vary along two dimensions (interpreted as "ability" and "preference for consumption"), which are behaviorally equivalent but normatively distinct. They work directly in a welfarist framework—maximizing a Pareto-weighted sum of utilities—which thus entails an assumption about how to make interpersonal comparisons; their parameterization is chosen so that whenever the population varies only on the preference dimension, the optimal tax policy is laissez-faire. (Dworczak ⓡ Kominers ⓡ Akbarpour (2021) take a similar modeling approach in a market design context, though without reference to the equal opportunity literature.)

As noted above, there are multiple approaches to conceptualizing equality of opportunity in the economics literature. This paper focuses on one particular version for the sake of concreteness, and specifically chooses to follow Roemer (1998) because it has inspired numerous applications, both theoretical and empirical (see e.g. Roemer and Trannoy (2016) for a survey). However, the main results generalize to other formulations, as discussed more in Section 6.

One alternative literature is that on "libertarian egalitarianism" based in the work of Fleurbaey (1994, 1995) (see Fleurbaey (2008) for a book-length treatment). This literature, like the approach here, assumes that agents differ in two kinds of characteristics, only one of which calls for compensation. However, there are two main conceptual differences. The first is that the "utilitarian" reward principle articulated here is replaced with a "natural" reward principle, which specifies that when agents differ only in e, all agents should receive the same amount of the good. The second is that the framework used is axiomatic, rather than based on a optimizing a numerical objective. (This makes it less obviously suited to comparing non-optimal allocations, which may be why this approach has received less attention outside of the theoretical literature.)[3]

A separate literature, following Kranich (1996) and Ok (1997), conceives of "opportunities" as elements of some abstract set and attempts to develop a theory of opportunity inequality analogous to the theory of income inequality. This theory is more challenging to apply to specific policy problems and, as far as this author is aware, has not informed empirical work.

The one paper that shares the same basic goal as this one—of understanding whether equal opportunity is equivalent to welfarism (and why or why not)—is Maniquet (2004). That paper obtained the surprising finding that equal opportunity can *always* be represented as welfarism. In Section 6 we return to discuss the relation between the two papers and account for their diverging results.

2 Preliminaries

We will write $\mathbb{R}_+$ for the set of nonnegative real numbers and $\mathbb{R}_{++}$ for the positive reals. Primes following names of single-variable functions (such as $H'(u)$) denote derivatives

[3]Also in contrast to Roemer (1993) and Van de gaer (1993), the strategy taken by Fleurbaey (1995) to deal with the tension between the reward and compensation principles is to relax one or the other by requiring it only in specific cases. This leads to the "egalitarian-equivalent" and "conditional equality" solutions in Fleurbaey (1995). Versions of these solutions can be formulated in our framework; both of them aim to equalize some measure of individual well-being across the population, so they are easily seen to be replicable by a welfarist criterion with the min aggregator.

unless otherwise indicated; primes elsewhere simply distinguish related variables (such as x and x').

Agents will be indexed by a pair (c, e) of characteristics (termed "circumstances" and "efforts" in the literature). We take as given open intervals $\overline{C} = (\underline{c}, \overline{c}) \subseteq \mathbb{R}$ and $\overline{E} = (\underline{e}, \overline{e}) \subseteq \mathbb{R}$, from which the values of (c, e) are drawn. The openness assumption will be convenient for many of our arguments, but it is not a substantive restriction here.

A *population* is a set of the form $P = C \times E$, where C, E are finite, nonempty subsets of $\overline{C}$ and $\overline{E}$ respectively. Note that this definition is restrictive: it imposes that c and e are independently distributed in the population, and also that there are no two agents at the same (c, e) pair. These restrictions are commonly imposed in previous literature, often for expository convenience.[4] In particular, the interpretation of a criterion such as (1.1) becomes more complicated when general populations are allowed. The restriction will not be problematic for us: our negative results on existence of welfarist representations will clearly continue to hold if more general populations are allowed, and in the cases where a welfarist representation does exist, it naturally extends to pick out a particular allocation in non-product populations.

A population is *one-dimensional* if $|C| = 1$ or $|E| = 1$, and *two-dimensional* otherwise. It is a *two-agent* population if $|C \times E| = 2$; two-agent populations are one-dimensional.

Given a population $P = C \times E$, an *allocation* on P is a function $X : P \to \mathbb{R}_+$, specifying how much of the good each agent in the population receives.

It will sometimes be convenient to notate an allocation X as a list $\left(\overset{c_1,e_1}{x_1}, \overset{c_2,e_2}{x_2}, \ldots, \overset{c_n,e_n}{x_n}\right)$, meaning that X is defined on a population whose elements are $(c_1, e_1), \ldots, (c_n, e_n)$, and $X(c_i, e_i) = x_i$.

A *distribution problem* consists of a pair $(P, \overline{x})$, where P is a population and $\overline{x} \in \mathbb{R}_+$. We may also refer to this as a *distribution problem on* P. Write $\Delta_P(\overline{x})$ for the set of allocations $X : P \to \mathbb{R}_+$ such that $\sum_{(c,e) \in P} X(c, e) = \overline{x}$. The interpretation of the problem $(P, \overline{x})$ is that the planner needs to choose an allocation in $\Delta_P(\overline{x})$.

Next, given a population P, we consider the planner's preference over allocations on it. A *preference*, notated $\succsim$, is a weak ordering (reflexive, transitive, complete) over

[4] In Roemer (1993, 1998, 2002), independence of c and e holds by *definition*: the analyst observes some proxy variable that may conflate effort and circumstances, and the relevant measure of effort is defined to be the quantile of this variable conditional on circumstances, which is therefore necessarily independent of c. That specific approach to defining e will not work here, as it is important that we allow variable populations with different choices of E (see Proposition 2.1). On the other hand, Roemer (2002) argues that his definition of e is appropriate because otherwise the distribution of effort is a property of the circumstance type, which, by hypothesis, agents should not be held responsible for; that defense for assuming independence between c and e does still apply here.

the allocations on P. We denote the asymmetric part of the preference by $\succ$ and the symmetric part by $\sim$. The preference is *monotone* if for all allocations X, X', the following are satisfied:

- if $X'(c,e) \geq X(c,e)$ for all $(c,e) \in P$, then $X' \succsim X$; and
- if $X'(c,e) > X(c,e)$ for all $(c,e) \in P$, then $X' \succ X$.

The preference is *continuous* if, for all X, the sets of allocations $\{X' \mid X' \succsim X\}$ and $\{X' \mid X \succsim X'\}$ are closed (in the natural topology).

To define equal-opportunity preferences, we take as exogenously given an *advantage* function $A : \overline{C} \times \overline{E} \times \mathbb{R}_+ \to \mathbb{R}$. We denote its typical arguments as (c, e, x). We assume that A is twice continuously differentiable in all variables jointly, and its derivative with respect to x is positive. (The examples (1.3), (1.4), (1.6) satisfy these assumptions, assuming all values of c and e are positive. The example (1.5) fails differentiability at $x = 0$, though this failure is irrelevant to the purpose of the example; it can also be avoided by replacing x with $x + \alpha$ for any constant $\alpha > 0$.)

An opportunitarian planner evaluates any allocation X on any population $P = C \times E$ using a numerical criterion $\mathcal{V}^{EOp}(X)$. An allocation X that maximizes $\mathcal{V}^{EOp}$ over $\Delta_P(\overline{x})$ is called an *equal-opportunity choice* for the distribution problem $(P, \overline{x})$. When P is one-dimensional, $\mathcal{V}^{EOp}$ is defined as follows:

- If $|E| = 1$, then $\mathcal{V}^{EOp}(X) = \min_{(c,e)\in P} A(c, e, X(c, e))$.
- If $|C| = 1$, then $\mathcal{V}^{EOp}(X) = \sum_{(c,e)\in P} A(c, e, X(c, e))$.

For now, we avoid specifying $\mathcal{V}^{EOp}$ when P is two-dimensional.

We now consider welfarist criteria to evaluate allocations. A *welfarist criterion* is a pair of functions of the form (U, W), where

- $U : \overline{C} \times \overline{E} \times \mathbb{R}_+ \to \mathbb{R}$, strictly increasing in its third argument x;
- $W : \cup_{n\geq 2} \mathbb{R}^n \to \mathbb{R}$, and for each n, the restriction of W to $\mathbb{R}^n$ is weakly increasing in all arguments and is symmetric (i.e. invariant to permutations of its arguments).

We interpret U as the measure of individual welfare, and W as the aggregator. In much of the paper we will only need to concern ourselves with the behavior of W on $\mathbb{R}^2$, but we define welfarist criteria for general population sizes to avoid multiple redundant definitions. For completeness we can also include $\mathbb{R}^1$ in the domain of W, and take W to be the identity there.

Note that we require W only to be weakly increasing, not strictly; this is natural given that the opportunitarian criterion uses the min operator.

We will often use the variable $\mathcal{W}$ to denote a welfarist criterion. We can naturally use such a criterion $\mathcal{W} = (U, W)$ to evaluate an allocation (on any population) by defining

$$\mathcal{W}\begin{pmatrix} c_1,e_1 & & c_n,e_n \\ x_1 & ,\ldots, & x_n \end{pmatrix} = W(U(c_1, e_1, x_1), \ldots, U(c_n, e_n, x_n)). \tag{2.1}$$

If the allocation is denoted by X, then we can write this value more simply as $\mathcal{W}(X)$. By symmetry of W, the order in which the agents are listed is irrelevant.

We say that the welfarist criterion $\mathcal{W}$ *is a welfarist representation of equal opportunity for one-dimensional populations*, or more simply that $\mathcal{W}$ *represents equal opportunity for one-dimensional populations*, if

$$\arg\max_{X \in \Delta_P(\overline{x})} \mathcal{V}^{EOp}(X) = \arg\max_{X \in \Delta_P(\overline{x})} \mathcal{W}(X) \tag{2.2}$$

for every distribution problem $(P, \overline{x})$ in which P is one-dimensional. Likewise, we will say that $\mathcal{W}$ *represents equal opportunity for two agents* if (2.2) holds whenever P is a two-agent population. We may say informally that $\mathcal{W}$ *represents equal opportunity* when we do not need to be precise about which version is meant. We are interested in understanding when a welfarist representation of equal opportunity exists.

Note that (2.2) requires exact equality of the sets of optimal choices for the two criteria. A weaker definition of representation would only require the maximizers of $\mathcal{V}^{EOp}$ to form a subset of the maximizers of $\mathcal{W}$, but then a welfarist representation would trivially always exist: we could take $U(c, e, x) = x$ and W to be the sum function, and then every allocation would be a maximizer of $\mathcal{W}$.

It may be helpful to briefly discuss what welfarism "means" in this context. Welfarism essentially requires anonymity at the aggregation stage. It does not mean symmetry with respect to the values of x; that is, it does not require $\mathcal{W}\begin{pmatrix} c_1,e_1 & c_2,e_2 \\ x_1 & x_2 \end{pmatrix} = \mathcal{W}\begin{pmatrix} c_1,e_1 & c_2,e_2 \\ x_2 & x_1 \end{pmatrix}$. Nor does it mean symmetry with respect to the individual advantage levels $A(c, e, x)$. It allows an individual well-being measure that has been adjusted to take account of the individual's attributes before being fed into the aggregator; what is required is that this adjustment should be independent of who else is in the population, or of how much of the good they receive. (The term "welfarism" is used in different ways in the social choice literature, but this usage seems to be consistent with the closest work, e.g. Maniquet (2004); Roemer (2002, pp. 460–461).)

We now give a couple of background results that will help motivate the specific formulation of our main results in the following sections. (Their proofs, and all others, are in the appendix.)

First, if we consider a single, *fixed* population P, then welfarism has no bite: *any* well-behaved preference over allocations that the planner could have can be represented by a welfarist criterion.

Proposition 2.1. *Consider any population $P = C \times E$, and any continuous, monotone preference $\succsim$ over allocations on P. There exists a welfarist criterion $\mathcal{W}$ such that, for all allocations X, X' on P,*

$$\mathcal{W}(X) \geq \mathcal{W}(X') \qquad \textit{if and only if} \qquad X \succsim X'.$$

The reason is that, when the population is fixed, the only restriction imposed by welfarism as we have defined it (aside from monotonicity) is that W should be symmetric; but the individual welfare function U can be chosen so that different agents' welfare have disjoint ranges, and then the symmetry requirement becomes vacuous.

This motivates instead allowing P to vary and requiring representation for all possible P, which is why we consider C and E in ambient ranges $\overline{C}, \overline{E}$. In order for comparisons across populations to have force, we need to further require U to be well-behaved with respect to changes in the individual characteristics. In particular, a continuity requirement seems natural. (Such a requirement then avoids the phenomenon in Proposition 2.1, because the ranges of $U(c, e, \cdot)$ cannot be disjoint for agents who are sufficiently close in (c, e)-space.)

We will actually consider the following stronger requirement. Say that $\mathcal{W} = (U, W)$ is *regular* if

- U is jointly continuously differentiable in its arguments (c, e, x), and $\partial U/\partial x > 0$ whenever $x > 0$; and
- for each n, W is continuous on $\mathbb{R}^n$.

It might seem inappropriate to require differentiability here, given that the opportunitarian uses a min criterion, which is not differentiable. However, note that we are requiring differentiability only on the individual welfare measure U, not the aggregator W—and we already assumed that A was differentiable. Thus, requiring regularity does not make the welfarist more constrained than the opportunitarian. (The author conjectures that Theorems 3.1, 4.4 will continue to hold even without this requirement.)

Our second background result says that, once the population can vary (and regularity is imposed), no welfarist criterion can agree with the equal-opportunity preferences over *all* allocations, even just for two-agent populations.

Proposition 2.2. *There is no regular welfarist criterion $\mathcal{W} = (U, W)$ such that, for every two-agent population $P = C \times E$ and any two allocations X, X' on P,*

$$\mathcal{W}(X) \geq \mathcal{W}(X') \qquad \textit{if and only if} \qquad \mathcal{V}^{EOp}(X) \geq \mathcal{V}^{EOp}(X').$$

The reason is quite simple: considering populations that differ only in c, $\mathcal{W}$ would have to replicate the min-advantage criterion and so be indifferent to increases in one agent's quantity (at least over the relevant part of its domain), whereas considering populations that differ only in e, $\mathcal{W}$ would need to be strictly increasing; we cannot have both of these.

For this reason, we focus on just seeking a $\mathcal{W}$ that agrees with $\mathcal{V}^{EOp}$ as to the optimal choices in distribution problems, rather than agreeing over all possible allocations. This goal requires nothing in terms of preferences between two non-optimal allocations of a given total quantity, nor between two allocations with different total quantities.

It bears mention that, while the preceding simple explanation for Proposition 2.2 relies on the "extremeness" of the min operator—its failure to be strictly monotone—the result itself does not. In fact, one can prove a similar impossibility result for any formulation of equal opportunity that entails more inequality aversion for agents differing in c than in e. The details will be left to a separate paper.

3 Linear advantage

We now proceed to study when a welfarist representation of equal opportunity exists. We begin with the advantage specification (1.3) from the opening example, which we restate here for convenience:

$$A(c, e, x) = (x + c)e.$$

We assume $\overline{C}, \overline{E} \subseteq \mathbb{R}_{++}$.

What would the equal-opportunity planner do, in any distribution problem with a one-dimensional population? If $|E| = 1$, so the population varies only in c, then the planner wishes to equalize A, which is done by giving each agent enough of the good so that $x + c$ is equalized in the population. If doing this requires more than the total available amount $\overline{x}$, then some agents (those with high c's) will simply be given 0, while

$x + c$ will be equalized among the remaining agents. If $|C| = 1$, so the population varies only in e, then the planner wishes to maximize total A, which entails giving all of the good to the one agent with the highest e.

Our first main result states that no regular welfarist criterion agrees with these choices—even if we only consider two-agent populations.

Theorem 3.1. *Assume that A is given by specification (1.3). There is no regular welfarist criterion that represents equal opportunity for two agents.*

An intuition is that the equal-opportunity choices involve a discontinuity as the two agents converge to the same (c, e) pair: if they are arbitrary close and differ only in c, the good is equally split, while if instead they differ only in e, the good is given entirely to the higher-e agent. Thus the first step toward Theorem 3.1 is:

Lemma 3.2. *Assume that A is given by (1.3). Suppose the regular welfarist criterion $\mathcal{W} = (U, W)$ represents equal opportunity for two agents. Then, for any $(c, e) \in \overline{C} \times \overline{E}$ and for any $x \in \mathbb{R}_{++}$ and $y \in [0, 2x]$,*

$$\mathcal{W}\left(\overset{c,e}{x}, \overset{c,e}{x}\right) = \mathcal{W}\left(\overset{c,e}{2x}, \overset{c,e}{0}\right) \geq \mathcal{W}\left(\overset{c,e}{2x - y}, \overset{c,e}{y}\right).$$

Noe that the $\mathcal{W}(\cdots)$ notation in the lemma is not, strictly speaking, the welfarist value of an allocation, since two agents with the same (c, e) pair do not constitute a population. (Indeed, if we had defined populations in a way that allows this, then Lemma 3.2 would point to an even more basic problem—the equal-opportunity choice is not defined for such a population.) Nonetheless we can still interpret this notation via (2.1).

This lemma alone is not enough for a contradiction; many welfarist criteria do satisfy its conclusion. Instead, the key step is the lemma below, which is not specific to the advantage specification (1.3) and will also be useful in the next section.

Suppose $[\underline{u}, \overline{u}] \subseteq \mathbb{R}$ is some interval, and $m > 0$. If the functions $H : [\underline{u}, \overline{u}] \to [0, m]$, $G : [0, 2m] \to \mathbb{R}$, and $\widetilde{W} : \mathbb{R}^2 \to \mathbb{R}$ satisfy

$$\widetilde{W}(u_1, u_2) = G(H(u_1) + H(u_2))$$

for all $u_1, u_2 \in [\underline{u}, \overline{u}]$, we call the pair (G, H) an *additive representation* for $\widetilde{W}$ on the interval $[\underline{u}, \overline{u}]$.

Lemma 3.3. *Let $\mu : [0,1] \to \overline{C} \times \overline{E} \times \mathbb{R}_{++}$ be a continuous curve; denote its component functions as $\mu(t) = (c(t), e(t), x(t))$. Let $\mathcal{W} = (U, W)$ be a regular welfarist criterion. Suppose that*

1. *the function $t \mapsto U(\mu(t))$ is a bijection from $[0,1]$ to a non-degenerate interval $[\underline{u}, \overline{u}] \subseteq \mathbb{R}$;*
2. *for every $t, t' \in [0,1]$, there exists $\overline{\epsilon} > 0$ such that*

$$\mathcal{W}\begin{pmatrix} {\scriptstyle c(t),e(t)} & {\scriptstyle c(t'),e(t')} \\ x(t) & , \ x(t') \end{pmatrix} \geq \mathcal{W}\begin{pmatrix} {\scriptstyle c(t),e(t)} & {\scriptstyle c(t'),e(t')} \\ x(t)+\epsilon, & x(t')-\epsilon \end{pmatrix}$$

for all $\epsilon \in (-\overline{\epsilon}, \overline{\epsilon})$, with strict inequality if $t \neq t'$ and $\epsilon \neq 0$.

Then, the restriction of W to $\mathbb{R}^2$ has an additive representation (G, H) on the interval $[\underline{u}, \overline{u}]$. Moreover, G, H are strictly increasing, and H is differentiable.

The essential meaning of the lemma is the following. Suppose that we have a curve in (c, e, x)-space such that the equal-opportunity planner, facing any distribution problem with two agents, would divide x between them so that they both land on the curve if possible. (Condition 2 of the lemma is simply the local optimality condition for the planner's problem.) If U takes distinct values along the path, the lemma tells us that W must have an additive representation on the corresponding interval of U-values. That is, roughly speaking, one of the following must be true: either U is constant along the path, or else W looks locally like the sum function, after an appropriate reparameterization of the individual utility measure.

Why is the lemma true? It essentially follows from the first-order condition for the equal-opportunity planner's problem. For a sketch, assume that both U and W are differentiable. Then, for any t and t', considering condition 2 of the lemma and using the first-order condition at $\epsilon = 0$ gives

$$\frac{d}{d\epsilon} W(U(c(t), e(t), x(t)+\epsilon), U(c(t'), e(t'), x(t')-\epsilon))|_{\epsilon=0} = 0$$

or, more explicitly,

$$\left.\frac{\partial W}{\partial u_1}\right|_{(u(t),u(t'))} \times \left.\frac{\partial U}{\partial x}\right|_{\mu(t)} - \left.\frac{\partial W}{\partial u_2}\right|_{(u(t),u(t'))} \times \left.\frac{\partial U}{\partial x}\right|_{\mu(t')} = 0 \tag{3.1}$$

(where arguments to W are being notated u_1, u_2, and $u(t), u(t')$ denote the values of U at points $\mu(t), \mu(t')$).

Rearranging,

$$\frac{\left.\frac{\partial W}{\partial u_1}\right|_{(u(t),u(t'))}}{\left.\frac{\partial W}{\partial u_2}\right|_{(u(t),u(t'))}} = \frac{\left.\frac{\partial U}{\partial x}\right|_{\mu(t')}}{\left.\frac{\partial U}{\partial x}\right|_{\mu(t)}}. \tag{3.2}$$

Suppose we define H by the differential equation

$$H'(u(t)) \times \left.\frac{\partial U}{\partial x}\right|_{\mu(t)} = 1.$$

Then (3.2) becomes

$$\frac{\left.\frac{\partial W}{\partial u_1}\right|_{(u(t),u(t'))}}{\left.\frac{\partial W}{\partial u_2}\right|_{(u(t),u(t'))}} = \frac{H'(u(t))}{H'(u(t'))}.$$

This says that, at the point $(u(t), u(t'))$, the level curves of the functions $W(u_1, u_2)$ and $H(u_1) + H(u_2)$ have the same slope. By varying t and t', this holds for all (u_1, u_2) in a neighborhood, thus the level curves must coincide, which means $W(u_1, u_2)$ can be expressed as a function of $H(u_1) + H(u_2)$.

The full proof of Lemma 3.3 shows that a more careful version of this argument works even without assuming that W is differentiable.

Once this lemma is in hand, completing the proof of Theorem 3.1 is relatively mechanical. For populations of two agents with the same e, the equal-opportunity planner would like to equalize their values of $x + c$—that is, to put them on a curve of the form

$$\{(c, e, x) \mid e \text{ constant}, \ x + c \text{ constant}\} \tag{3.3}$$

if possible. If the value of U along such a curve is non-constant, we can apply Lemma 3.3 to obtain an additive representation for W over an interval. A few calculations then show that it is impossible for such a $\mathcal{W}$ to satisfy the equality in Lemma 3.2—that is, given a "population" of two agents at the same (c, e), to be indifferent between allocating the good equally and allocating all to one agent—while strictly preferring the equalizing allocation for two agents at the same e but distinct nearby c's, as the equal-opportunity planner should. This leaves only the possibility that the value of U is constant along each curve as in (3.3), but this possibility is easy to rule out directly.

4 Concave advantage

We now get rid of the discontinuity in equal-opportunity allocations that was a driving force behind Theorem 3.1 by assuming that the advantage function A is strictly concave in x. More specifically, we will rely on the following assumption throughout the remaining results.

Assumption 4.1. *The function $A(c, e, x)$ satisfies the following everywhere:*

- $\partial^2 A/\partial x^2 < 0$.
- *A is weakly increasing in c, and $\partial A/\partial x$ is weakly increasing in e.*

The role played by the monotonicity assumptions with respect to c and e will be explained later.

(We could alternatively replace either instance of "increasing" by "decreasing" in the second part of Assumption 4.1; this is tantamount to reversing the order on c or e. However, the version here seems most natural since it implies that the equal-opportunity planner would be inclined to give more of the resource to lower-c and higher-e agents.)

For some results it will be useful to have the following slightly stronger assumption, which allows us to extend all the differentiability conditions to a neighborhood of the $x = 0$ boundary:

Assumption 4.2. *There exist a function $\widetilde{A} : \overline{C} \times \overline{E} \times \mathbb{R}_+ \to \mathbb{R}$ and a constant $\alpha > 0$ such that $A(c, e, x) = \widetilde{A}(c, e, x + \alpha)$ for all (c, e, x), and $\widetilde{A}$ is twice continuously differentiable, satisfies $\partial\widetilde{A}/\partial x > 0$, and satisfies all the conditions of Assumption 4.1.*

To comment briefly on the examples from the introduction in relation to these assumptions: specification (1.4) satisfies Assumption 4.2 as long as values of c, e are bounded strictly above 0; (1.5) can be tweaked to satisfy it (as before) by replacing x by $x + \alpha$ for any positive α; and (1.6) satisfies it if c, e are bounded above 0, x is replaced by $x + \alpha$ where α is a constant that depends on $\overline{C}$ and $\overline{E}$, and the ordering on e is reversed (see the proof of Proposition 5.7).

It will be useful to write $B = -\partial A/\partial x$. The function $B(c, e, x)$ is negative-valued and strictly increasing in x. B plays a role symmetric to A from the point of view of the equal-opportunity planner: for a population with constant e but different c's, the planner wishes to allocate the good to equalize A; for a population with constant c but different e's, the planner wishes to equalize B (since this corresponds to the first-order condition to maximize the sum of A across agents).

More formally, we can describe the equal-opportunity choices as follows.

Lemma 4.3. *Let $P = C \times E$ be a one-dimensional population. For any distribution problem $(P, \overline{x})$, there is a unique allocation X^* that maximizes $\mathcal{V}^{EOp}$, and it can be characterized as follows:*

1. *If $|E| = 1$, then X^* is the unique allocation of total quantity $\overline{x}$ having the following property: there is a value A^* such that every agent (c, e) with $A(c, e, 0) \leq A^*$ is given a quantity x such that $A(c, e, x) = A^*$, and every agent with $A(c, e, 0) > A^*$ is given* 0.

2. *If $|C| = 1$, then X^* is the unique allocation of total quantity $\overline{x}$ having the following property: there is a value B^* such that every agent (c, e) with $B(c, e, 0) \leq B^*$ is given a quantity x such that $B(c, e, x) = B^*$, and every agent with $B(c, e, 0) > B^*$ is given* 0.

It is useful to conceptualize these optimal choices in terms of curves in (c, e, x)-space. Holding fixed the value of e, and varying c and x so as to keep A constant, we trace out a curve, which we call an *A-curve*. That is, an A-curve is a set of the form

$$\{(c, e, x) \in \overline{C} \times \overline{E} \times \mathbb{R}_+ \mid e = e^*, A(c, e, x) = A^*\}$$

for some e^* and A^*. Likewise, we define a B-curve to be a set of the form

$$\{(c, e, x) \in \overline{C} \times \overline{E} \times \mathbb{R}_+ \mid c = c^*, B(c, e, x) = B^*\}$$

for some c^* and B^*. Note that the monotonicity conditions in Assumption 4.1 ensure that each A-curve is connected, with x being a (weakly) decreasing function of c along the curve; and each B-curve is connected, with x being a (weakly) increasing function of e along the curve.

In any population of agents with the same e, the equal-opportunity planner wishes to put all agents on the same A-curve (aside from those who get zero and are still above this curve); in any population of agents with the same c, the planner wishes to put them all on the same B-curve. These properties are what we use below.

The key condition to have a welfarist representation of equal opportunity is that (c, e, x)-space can be divided into surfaces, with each surface expressible as a union of A-curves and also as a union of B-curves. Formally: We define an *AB-foliation* to be an equivalence relation $\approx$ on the set $\overline{C} \times \overline{E} \times \mathbb{R}_{++}$, such that for all $(c, e, x), (c', e', x') \in \overline{C} \times \overline{E} \times \mathbb{R}_{++}$:

- if $e = e'$, then $(c, e, x) \approx (c', e', x')$ if and only if $A(c, e, x) = A(c', e', x')$;
- if $c = c'$, then $(c, e, x) \approx (c', e', x')$ if and only if $B(c, e, x) = B(c', e', x')$.

(Note that we have restricted to $x > 0$; it will be convenient to work with this open set and then deal with $x = 0$ later.) The equivalence classes are called *leaves* of the foliation.

Intuitively, an AB-foliation exists if, whenever we start at a point (c, e, x) and walk through space along A-curves and B-curves, if we return to the initial values of (c, e), then we also return to the initial value of x.

With this definition, we can state our main results.

Theorem 4.4. *Suppose that Assumption 4.1 holds. There is no regular welfarist criterion that represents equal opportunity for two agents unless an AB-foliation exists.*

(Thus, when the foliation does not exist, this theorem—like Theorem 3.1—says that a welfarist cannot match the equal-opportunity choices even if we only consider two-agent populations.)

Theorem 4.5. *Suppose that Assumption 4.1 holds, and an AB-foliation exists. Then there is a welfarist criterion (U, W) that represents equal opportunity for one-dimensional populations. In fact, there exists such a criterion in which the aggregator W is the min operator; there is also one in which W is the sum operator. Moreover, if Assumption 4.2 holds, these welfarist criteria are regular.*

The intuition behind Theorem 4.5 is relatively straightforward. If an AB-foliation exists, we can index the leaves from "lowest" to "highest," and then define $U(c, e, x)$ to be the index of the leaf in which point (c, e, x) lies. This U, together with the min operator for W, constitutes a welfarist representation of equal opportunity (for any one-dimensional population): indeed, this welfarist criterion is maximized by putting all agents on the same leaf, to the extent possible, which corresponds in one-dimensional populations to putting them all on the same A-curve or B-curve. Moreover, to generate a welfarist representation $(\widetilde{U}, \widetilde{W})$, where $\widetilde{W}$ is the sum function, notice that allocating x to maximize the sum of $\widetilde{U}$ corresponds to equalizing the marginal values $\partial\widetilde{U}/\partial x$, so it suffices to define $\widetilde{U}$ in such a way that $\partial\widetilde{U}/\partial x = Z(U)$ where Z is a decreasing function (in order to ensure the maximand is concave).

Theorem 4.4 says that existence of an AB-foliation is also a necessary condition for a regular welfarist representation. The reason for this is a bit more subtle. If we cannot construct an AB-foliation, this means that we can walk a path through (c, e, x)-space along A-curves and B-curves and return to our starting (c, e)-coordinates, but at a different value of x (see Figure 2(a)). This already precludes a welfarist representation in which W is the min function: such a representation would mean that the planner wants to equalize values of U, so U should be constant on each A-curve and B-curve, and therefore constant along the whole path, but this is impossible because U should be strictly increasing in x. A slight extension of this argument shows that we also cannot have a regular welfarist representation in which W is the sum function: Such a representation would require $\partial U/\partial x$ to be constant along the path, and then by perturbing the path (as in Figure 2(b)), we can reach a region of other points in (c, e, x)-space, so $\partial U/\partial x$ must be constant throughout this region; but then all allocations contained in this region would give equal welfare, contradicting that the equal-opportunity allocation should be the unique maximizer.

To build on this reasoning to rule out *any* aggregator function, we use Lemma 3.3. Because U cannot be constant throughout the path (as we already saw), the lemma applies, and it implies the aggregator can be written as $W(u_1, u_2) = G(H(u_1) + H(u_2))$, at least locally. Then maximizing W is equivalent to just maximizing $H(u_1) + H(u_2)$, and so we have a regular welfarist representation where the individual utility measure is $H \circ U$ and the aggregator is the sum function—but we already argued for why the latter property is impossible.

To summarize, if a regular welfarist representation exists, either the utility measure U is constant along A-curves and B-curves, in which case we can use it to construct the foliation; or else we can apply Lemma 3.3 to obtain an additive representation, and in such a representation, *marginal* utility is constant along A-curves and B-curves, which again gives us the foliation.

Some additional complexity in the proofs arises from the fact that Lemma 3.3 applies only locally, and so in the proof of Theorem 4.4, we must show that the foliation exists locally—in a neighborhood of each (c, e, x) point—and then patch together the local pieces. It is here that the second part of Assumption 4.1 plays a crucial role: it ensures that each A-curve and B-curve is connected, so that gluing together local pieces does give us an entire foliation. The assumption also plays a role in the proof of Theorem 4.5, because each foliation leaf may cover only a portion of the rectangle $\overline{C} \times \overline{E}$ (that is, as we move along an A-curve, x may fall to zero or shoot up to infinity without reaching

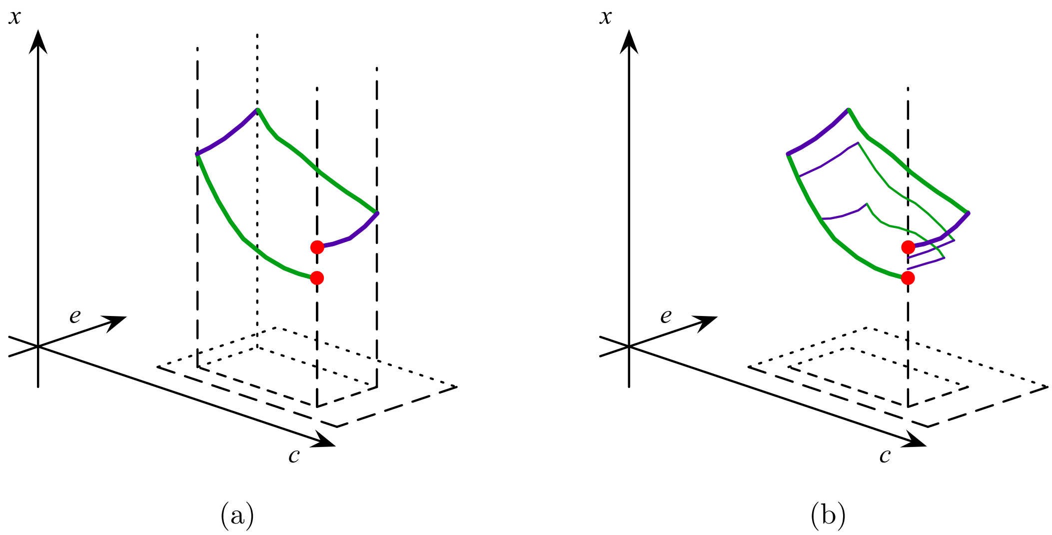

Figure 2: (a) A non-closing path along A-curves and B-curves. (b) Perturbing the path.

all c-values, and likewise along B-curves). The monotonicity assumption ensures that even when different leaves cover disjoint parts of this rectangle, we have an unambiguous notion of which leaf is higher than the other.

We can illustrate Theorem 4.4 by verifying that, for the example advantage specification (1.5) from the introduction, there is no AB-foliation, and therefore the theorem implies that no regular welfarist criterion can represent equal opportunity. We have not specified $\overline{C}, \overline{E}$ for this example, but they can be arbitrary intervals in $\mathbb{R}_{++}$. Take $c, c' \in \overline{C}$ and $e, e' \in \overline{E}$ with $c < c'$ and $e < e'$. If the foliation exists, then the following must all lie in the same leaf (adjacent points are either on the same A-curve or the same B-curve):

$$(c, e, 1) \approx \left(c, e', \frac{c+e'}{c+e}\right) \approx \left(c', e', \frac{(c+e')^2}{(c+e)(c'+e')}\right)$$

$$\approx \left(c', e, \frac{(c+e')^2(c'+e)}{(c+e)(c'+e')^2}\right) \approx \left(c, e, \frac{(c+e')^2(c'+e)^{2\prime}}{(c+e)^2(c'+e')^2}\right)$$

But since the fraction in the third coordinate of the last point is greater than 1, this point cannot be in the same foliation leaf as $(c, e, 1)$, a contradiction.

We can also use the analysis to address another claim made in the introduction: that the advantage function (1.6) does not allow a regular welfarist representation of equal opportunity with $U = A$. We just sketch the argument. Because $U = A$ is not constant along B-curves, Lemma 3.3 implies that the welfare aggregator would have to have an additive representation (G, H), and so, as argued above, marginal utility in this representation, namely $\frac{\partial}{\partial x}(H \circ A)$, would have to be constant along A-curves and B-curves. But this marginal utility equals $H'(A) \times \partial A/\partial x$, and $H'(A)$ is automatically constant along A-curves while $\partial A/\partial x$ is not, so their product cannot be—a contradiction.

We can explore a bit further the class of advantage specifications for which an AB-foliation does exist; this discussion will be left somewhat informal. If the foliation exists, then we can consider a welfare function U that indexes the leaves as in Theorem 4.5. For any fixed value of e, there is a one-to-one (and increasing) correspondence between A-curves and values of U, and thus we have the equation

$$U = I_e(A)$$

for some family of increasing functions I_e, parameterized by e. Likewise, for any fixed value of c, we have the equation

$$U = J_c(B)$$

for some family of increasing functions J_c, parameterized by c. Consequently, for any given choice of the families $(I_e), (J_c)$, we can write the equation

$$B = J_c^{-1}(I_e(A)). \tag{4.1}$$

Now, for each fixed (c, e), we can use the fact that $B = -\partial A/\partial x$ (the first time we use this relation!) to write (4.1) explicitly as a homogeneous differential equation governing the function $A(c, e, \cdot)$:

$$\left.\frac{\partial A}{\partial x}\right|_{(c,e,x)} = -J_c^{-1}(I_e(A(c, e, x))). \tag{4.2}$$

For each (c, e) pair, let $K_{c,e}(t)$ be an antiderivative of the function $-1/J_c^{-1}(I_e(t))$, normalized by taking (say) $K_{c,e}(0) = 0$ for each (c, e), and write $L_{c,e}$ for the inverse of $K_{c,e}$. Then we get

$$A(x) = L_{c,e}(x)$$

as a solution to the differential equation, or more generally

$$A(x) = L_{c,e}(x + \beta(c, e))$$

where $\beta(c, e)$ is any function of (c, e) (subject to the smoothness and monotonicity requirements).

This provides a recipe for generating many examples of advantage specifications that are consistent with an AB-foliation. For example, if we simply take I_e to be the identity function for each e, then K is parameterized by c alone, thus so is L, and we can write

$$A(x) = L_c(x + \beta(c, e)).$$

That is, e enters into the advantage function only through the additive shifter $\beta(c, e)$. And indeed, for any advantage function that can be decomposed in this way (with L_c strictly concave), the equal-opportunity choices simply correspond to equalizing A over the population—thus we have a welfarist representation where the individual welfare measure is A itself, and the aggregator is the min. Similarly, if we instead take J_c to be the identity function for each c, then L is parameterized by e alone, and we get

$$A(x) = L_e(x + \beta(c, e)).$$

For any advantage function of this form, the equal-opportunity choices correspond to minimizing the sum of A over the population—so we have a welfarist representation where the individual welfare measure is A and the aggregator is the sum (such as the example (1.4) from the introduction). If we take nontrivial choices for both I_e and J_c then we can generate more complex examples of A's. For example, by taking a linear specification, $I_e(A) = e(A - 1)$ and $J_c(B) = B/c$, then (4.2) reads $\partial A/\partial x = ce(1 - A)$, which is how the example (1.6) was generated.

How much flexibility in the choice of A is afforded by the condition (4.2)? Essentially, for any fixed c^* and e^*, if we specify the values of A at points of the form $(c^*, \cdot, \cdot)$ and $(\cdot, e^*, \cdot)$, then (4.2) pins down the rest of A, up to the choice of the shifters $\beta(c, e)$. We can see this because if we write $M_{c,e} = J_c^{-1} \circ I_e$, then (4.2) identifies the function $M_{c^*,e}$ for all e and M_{c,e^*} for all c, and then these determine the function $M_{c,e}$ for any other (c, e)-pair via the relation

$$M_{c,e} = M_{c,e^*} \circ M_{c^*,e^*}^{-1} \circ M_{c^*,e}$$

and thus $L_{c,e}$ is also determined for each (c, e)-pair. This is a way to see that the class of specifications of A that allows for an AB-foliation is rich but very far from generic.

5 Larger populations

So far, we have focused exclusively on one-dimensional populations. When an AB-foliation does not exist, a (regular) welfarist criterion cannot represent equal opportunity even for these special populations. However, when it does exist, such a representation is possible, and it is natural to ask whether the representation extends to general populations. This requires first specifying what the equal-opportunity choice is in general; recall that it can be defined in various ways—by maximizing objective (1.1), or (1.2), or something else.

We begin by pointing out that when an AB-foliation does exist, it already suggests a natural solution to any distribution problem: namely, give each agent an amount that puts them on the same foliation leaf (or zero, in corner cases).

Formally, assume that the advantage function A satisfies Assumption 4.1 and that an AB-foliation exists. Let U be the function that, together with the min aggregator, provides a welfarist representation of equal opportunity for one-dimensional populations (from Theorem 4.5). For any distribution problem $(P, \overline{x})$, define the *canonical allocation* as follows: for some value u^*, each agent $(c, e) \in P$ with $U(c, e, 0) < u^*$ receives x such that $U(c, e, x) = u^*$, and each agent (c, e) with $U(c, e, 0) \geq u^*$ receives 0; the value of u^* is determined such that the total quantity allocated is exactly $\overline{x}$. Note that this allocation indeed exists and is unique.[5]

Note also that the specific U constructed in the proof of Theorem 4.5 is not essential here; any U that indexes the leaves of the foliation will give rise to the same canonical allocation.

It is immediate that the welfarist criterion $(U, \min)$ selects the canonical allocation in any distribution problem. (The argument is similar to Lemma 4.3.) However, there can be other welfarist criteria that represent equal opportunity for one-dimensional populations but select a different allocation in some two-dimensional populations. This suggests that our canonical allocation may not be quite so canonical. This is because the requirement of representing equal opportunity for one-dimensional populations does not constrain the shape of the welfare aggregator at points $(u_1, \ldots, u_n)$ that can arise only in a two-

[5] To see existence, pick a value $\underline{u} < \min_{(c,e)\in P} U(c, e, 0)$ and $\overline{u} = \min_{(c,e)\in P} U(c, e, \overline{x})$. For each $u \in [\underline{u}, \overline{u}]$, define $x^{-1}(u|c, e)$ to be the quantity x such that $U(c, e, x) = u$, or 0 if $U(c, e, 0) \geq u$. This is a weakly increasing, continuous function of u. So, $\sum_{(c,e)\in P} x^{-1}(u|c, e)$ is also weakly increasing and continuous in u, running from 0 at $\underline{u}$ to a value $\geq \overline{x}$ at $\overline{u}$, so it hits $\overline{x}$ at some u in between, and this determines u^*. To see uniqueness, note that if two allocations X^*, X^{**} both satisfy the conditions for the canonical allocation with corresponding welfare levels $u^* < u^{**}$, then $X^*(c, e) \leq X^{**}(c, e)$ for all (c, e), with strict inequality for any agent who receives a positive amount at either allocation, so they cannot both sum to $\overline{x}$.

dimensional population.

For a concrete example, let $\overline{C} = \overline{E} = (0, 8)$, and suppose the advantage function is given by $A(c, e, x) = x/(1 + x)$. This satisfies the differentiability requirements and Assumption 4.1. It has an AB-foliation in which all sheets are flat; thus, in any distribution problem, the canonical solution would simply divide the good equally. Now consider the (regular) welfarist criterion $\mathcal{W} = (U, W)$, where

$$U(c, e, x) = \frac{x}{1 + x} + c + e;$$

$$W(u_1, \ldots, u_n) = \sum_i u_i + \epsilon \cdot \max\{0, \max_i u_i - \min_i u_i - 10\},$$

where ϵ is a small positive constant. Note that in any one-dimensional population, in any allocation X, the difference between the highest and lowest values of $U(c, e, X(c, e))$ is less than 10 (because the $x/(1 + x)$ terms are all between 0 and 1, and either all c terms are equal and the range of e-values is less than 8, or vice versa). Therefore, the second term in the definition of W will always be zero, and maximizing $\mathcal{W}$ is equivalent to maximizing the sum of U, which is achieved by dividing the good equally. Thus, this criterion represents equal opportunity for any one-dimensional population. But consider the two-dimensional population $P = \{1, 7\} \times \{1, 7\}$, and quantity 4 of the good. The canonical solution would divide the good equally. However, locally near this allocation, the welfarist criterion is linear in utility with a weight $1 + \epsilon$ on the $(7, 7)$ agent and weight $1 - \epsilon$ on the $(1, 1)$ agent, so this criterion would be improved by taking some of the good away from the $(1, 1)$ agent and giving it to the $(7, 7)$ agent.

On the other hand, the canonical allocation is indeed singled out if we further restrict ourselves to welfarist criteria in which the aggregator is separable, as in Debreu (1959); Gorman (1968). Essentially, these are the criteria for which the planner's preference over allocations of the good within a subset of agents does not depend on the quantities received by the remaining agents.

More specifically, say that a symmetric function $W : \cup_{n \geq 2} \mathbb{R}^n \to \mathbb{R}$ is *weakly separable* if, for all $1 \leq m < n$, all $(x_1, \ldots, x_n)$ and $(x'_1, \ldots, x'_n)$, if

$$W(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) > W(x'_1, \ldots, x'_m, x_{m+1}, \ldots, x_n),$$

then

$$W(x_1, \ldots, x_m, x'_{m+1}, \ldots, x'_n) \geq W(x'_1, \ldots, x'_m, x'_{m+1}, \ldots, x'_n).$$

The sum and min are familiar examples of weakly separable aggregators; many others, such as the median, are not. (The specific combination of strict and weak inequalities in the definition is needed in order for the min function to satisfy it.) Then:

Proposition 5.1. *Suppose the advantage function satisfies Assumption 4.2 and an AB-foliation exists. Then, in any distribution problem* $(P, \overline{x})$*, the canonical allocation* $X^*(P, \overline{x})$ *is the unique allocation* X *with the following property: for any regular welfarist criterion* $\mathcal{W} = (U, W)$ *that represents equal opportunity for one-dimensional populations and such that* W *is weakly separable,* X *maximizes* $\mathcal{W}$ *over* $\Delta_{\overline{x}}(P)$*.*

We take this henceforth as sufficient justification for focusing on the canonical allocation.

Now, how does this compare to the allocations selected by equal-opportunity criteria in the literature? As noted in the introduction, there are two main such criteria[6]: the "sum-of-mins" criterion (1.1), which we write out more explicitly as

$$\mathcal{V}^{\Sigma m}(X) = \sum_{e \in E} \left(\min_{c \in C} A(c, e, X(c, e)) \right)$$

(for any population $C \times E$ and any allocation X on it); and the "min-of-sums" criterion (1.2):

$$\mathcal{V}^{m\Sigma}(X) = \min_{c \in C} \left(\sum_{e \in E} A(c, e, X(c, e)) \right).$$

We note that these each select a unique allocation:

Lemma 5.2. *As long as the advantage function* A *is strictly concave in* x*, for each distribution problem* $(P, \overline{x})$*, there is a unique allocation that maximizes* $\mathcal{V}^{\Sigma m}$*, and also a unique allocation that maximizes* $\mathcal{V}^{m\Sigma}$*.*

We will accordingly refer to these allocations as $X^{\Sigma m}(P, \overline{x})$ and $X^{m\Sigma}(P, \overline{x})$ (or just $X^{\Sigma m}$ and $X^{m\Sigma}$ when the arguments are clear from context). We will also write $X^*(P, \overline{x})$ (or just X^*) for the canonical allocation.

In general, the canonical allocation, the sum-of-mins allocation, and the min-of-sums allocation need not coincide. In particular, they can differ in "corner" cases where some agents receive zero of the good.

[6] In the literature, e.g. Roemer and Trannoy (2016), these criteria have also been called the "mean-of-mins" and "min-of-means." The distinction between sum and mean is of course not substantive; by using sums, we avoid carrying around a $1/|E|$ multiplicative factor.

For a simple example, let $\overline{C} = \overline{E} = (1, 5)$, and suppose advantage is given by $A(c, e, x) = \sqrt{(x + c)e}$ as in (1.4). There is an AB-foliation; on each foliation leaf, $e/(x + c)$ is constant. Consider the population $C = E = \{2, 4\}$, and the total quantity $\overline{x} = 7$. The canonical solution X^* is the following allocation:

	$e = 2$	$e = 4$
$c = 2$	1	4
$c = 4$	0	2

This puts agent $(4, 2)$ at the corner and puts the other three agents on the same foliation leaf.

To see that $X^{\Sigma m}$ is different from X^*, it suffices to check that X^* does not maximize the sum-of-mins criterion. Consider starting from X^* and reducing the amount given to the $(2, 4)$ and $(4, 4)$ agents by ϵ each, and increasing the amount given to $(2, 2)$ by 2ϵ. The sum-of-mins criterion evaluates this allocation at

$$\sqrt{(3 + 2\epsilon) \cdot 2} + \sqrt{(6 - \epsilon) \cdot 4} = \sqrt{6 + 4\epsilon} + \sqrt{24 - 4\epsilon}$$

which is increasing in ϵ at $\epsilon = 0$. Thus a perturbation in this direction would be an improvement.

And improving the min-of-sums criterion is even easier: at the canonical solution, the sum of advantages for the $c = 2$ agents is strictly below that of the $c = 4$ agents, so taking some of the good from agent $(4, 4)$ and giving it to agent $(2, 4)$ would lead to an improvement.

On the other hand, this phenomenon is specific to corner solutions in the sense that it does not arise when all agents are allocated positive quantities of the good:

Proposition 5.3. *Suppose that A satisfies Assumption 4.1, and that an AB-foliation exists. Consider any distribution problem $(P, \overline{x})$. If at least one of the allocations $X^*(P, \overline{x})$, $X^{\Sigma m}(P, \overline{x})$, $X^{m\Sigma}(P, \overline{x})$ gives positive quantities to all agents, then all three of these allocations coincide.*

The result follows relatively easily from examining the marginal conditions to optimize the sum-of-mins and min-of-sums criteria.

In some cases, this equivalence also holds even when some agents are at corners. The key condition is that, for each leaf of the foliation, the set of (c, e) pairs that it covers has a product structure. When this happens, in the canonical allocation for any distribution problem, the set of agents receiving positive quantities also has a product structure, thereby avoiding the phenomena that drive the different solutions apart in the preceding example.

To formalize this, continue to suppose that Assumption 4.1 holds and an AB-foliation exists. We say that *the flat-curve condition is satisfied* if, for all $(c, e) \in \overline{C} \times \overline{E}$, either $A(c', e, 0) = A(c, e, 0)$ for all $c' < c$ or $B(c, e', 0) = B(c, e, 0)$ for all $e' > e$ (or both). Intuitively, this says that either the A-curve or B-curve through $(c, e, 0)$ is flat rather than extending to points above $x = 0$.

An equivalent condition is as follows. Let U again be the function that indexes the foliation leaves (constructed in the proof of Theorem 4.5). Then say that *the product-domain condition is satisfied* if, for every $u^* \in \mathbb{R}$, there exist thresholds $\widehat{c}(u^*) \in [\underline{c}, \overline{c}]$ and $\widehat{e}(u^*) \in [\underline{e}, \overline{e}]$ such that, for all c, e, we have $U(c, e, 0) < u^*$ if and only if $c < \widehat{c}(u^*)$ and $e > \widehat{e}(u^*)$. (Note that we can have $\widehat{c}(u^*) = \underline{c}$ or $\overline{c}$, in which case the requirement $c < \widehat{c}(u^*)$ is never satisfied or always satisfied, respectively, and similarly for $\widehat{e}(u^*)$. Note also that if this condition is satisfied for this particular function U, then it also holds for any other continuous U that indexes the foliation leaves.)

Lemma 5.4. *Suppose the advantage function satisfies Assumption 4.1 and an AB-foliation exists. The flat-curve condition is satisfied if and only if the product-domain condition is satisfied.*

Proposition 5.5. *Suppose the advantage function satisfies Assumption 4.1 and an AB-foliation exists. If the flat-curve condition is satisfied, then for every distribution problem $(P, \overline{x})$, the allocations $X^*(P, \overline{x}), X^{\Sigma m}(P, \overline{x})$, and $X^{m\Sigma}(P, \overline{x})$ coincide.*

Conversely, if the flat-curve condition is not satisfied, then there exists a distribution problem $(P, \overline{x})$, where $P = C \times E$ with $|C| = |E| = 2$, such that the three allocations $X^*(P, \overline{x}), X^{\Sigma m}(P, \overline{x}), X^{m\Sigma}(P, \overline{x})$ are all different.

Although we have been assuming existence of an AB-foliation, in fact, the second part of Proposition 5.5 holds even without assuming this. When such a foliation does not exist, the canonical allocation X^* is not defined in general. However, we do have:

Proposition 5.6. *Suppose that the advantage function satisfies Assumption 4.1, and for every distribution problem $(P, \overline{x})$, the allocations $X^{\Sigma m}(P, \overline{x})$ and $X^{m\Sigma}(P, \overline{x})$ coincide. Then, an AB-foliation exists.*

And in this case, the flat-curve condition must be satisfied (by the second part of Proposition 5.5), so in fact all three solutions coincide in every distribution problem.

Our next result is of a somewhat different flavor. We have seen that, even when there is an AB-foliation so that the canonical allocation is defined, the sum-of-mins and min-of-sums criteria may fail to select it. But perhaps this is just telling us that these criteria are not the right way of defining the equal-opportunity choices. Is there some other way to aggregate advantage that does always select the canonical allocation as optimal? Ideally, we might hope to find a unique such aggregator; if so, that would suggest we could use this aggregator to define the equal-opportunity criterion $\mathcal{V}^{EOp}$ more generally.

Unfortunately, the result below tells us that such an aggregator does not exist.

To state the result, we now need to allow the advantage function A to vary. Then the canonical allocation in any distribution problem depends on A; write $X^*_A(P, \overline{x})$ accordingly.

We also define an *advantage allocation* to be a set of triples, of the form

$$\{(c, e, y(c, e)) \mid c \in C, e \in E\}$$

for some population $C \times E$ and some function $y : C \times E \to \mathbb{R}$. For any specification of A and any allocation X over a population $C \times E$, we can define the resulting advantage allocation

$$A[X] = \{(c, e, A(c, e, X(c, e))) \mid (c, e) \in C \times E\}.$$

And now define an *advantage aggregator* $\mathcal{V}$ to be a function that maps every advantage allocation (over any population) to a real number. Note that, unlike welfare aggregators W that simply take as input a numerical list of welfare values, an advantage aggregator needs to be told which agent (c, e) is associated with each advantage value. This is natural, since the standard aggregators (such as the sum-of-mins and min-of-sums) make use of this information.

Proposition 5.7. *There is no advantage aggregator $\mathcal{V}$ that has the following property: For every specification of the advantage function A that satisfies Assumption 4.1 and has an AB-foliation, and every distribution problem $(P, \overline{x})$,*

$$\arg\max_{X \in \Delta_P(\overline{x})} \mathcal{V}(A[X]) = \{X^*_A(P, \overline{x})\}. \tag{5.1}$$

Moreover, this impossibility arises even if we only consider distribution problems in which $P = C \times E$ with $|C| = |E| = 2$.

Note that this proposition indeed requires allowing the specification of A to vary, unlike all of the other results in the paper. If we considered only a single, fixed A, then we could trivially satisfy (5.1) in all distribution problems, by just defining $\mathcal{V}(X)$ to be 1 whenever X is the canonical allocation in some problem and 0 otherwise.

6 Discussion

We have adopted a formulation of equal opportunity based on Roemer (1993, 1998) for the sake of concreteness. However, it should be clear that much of the analysis applies equally well to many other formulations with the same basic structure: two-dimensional heterogeneity among potential agents, one basic criterion used to evaluate allocations in constant-c populations, and a different criterion for constant-e populations. As long as the equal-opportunity allocations in one-dimensional populations are defined so as to satisfy Lemma 4.3, for some functions A and B that have appropriate smoothness and monotonicity properties, then versions of our main results, Theorems 4.4 and 4.5, will hold. Indeed, these results make no use of the fact that A and B are related by $B = -\partial A/\partial x$. Thus, for example, instead of adopting complete inequality-aversion and inequality-neutrality (the min and sum operators) for the Compensation and Reward principles, we could use some alternative, less extreme criteria. Alternatively, we could keep the extreme inequality attitudes but assume the planner has different individual-level objectives, say $\min A$ in constant-e populations and $\sum A'$ in constant-c populations, where A' is a function different from A. (This might be natural, say, in a situation of allocating scholarships: we could imagine a compensation principle that says scholarship funds should be allocated across c's to equalize the benefit that students derive, but allocated across e's to maximize total *social* benefits, which could differ from individual benefits.) Finally, we could replace our reward principle by a version of the principle used in Fleurbaey (1995), specifying that in constant-c populations the good should be divided equally: simply take $B(c, e, x) = x$.

This paper draws inspiration from Maniquet (2004), who also studies the possible equivalence between equal opportunity and welfarism. This paper adopts major features from that one, in particular the setup with variable sets of agents and the focus on distribution problems. However, Maniquet works in the libertarian egalitarianism framework of Fleurbaey (1994, 1995). In particular, the approach is axiomatic: Maniquet proposes a set of axioms for a social preference order to capture equality of opportunity, and shows that *any* preference satisfying these axioms can be represented as welfarist—a rather different conclusion than in this paper. Importantly, one of Maniquet's axioms is a consistency axiom (a close relative of separability), requiring that when two allocations differ only for a subset of agents, the planner's preference between them should be the same as if the agents outside this subset were not present. This is a significant restriction: for example, in our setting, the criteria (1.1) and (1.2) do not satisfy it. In fact, under the assumptions of Section 4, when there is an AB-foliation and thus a welfarist representation of equal opportunity exists, the sum representation does satisfy consistency. Conversely, when there is no AB-foliation, one can readily check that *no* social preference over allocations satisfying consistency will always select the equal-opportunity allocation in two-agent distribution problems (even without worrying about welfarism). This implies that, in our framework, the notion of equal opportunity satisfies consistency if and only if it has a regular welfarist representation. This offers a new perspective on Maniquet (2004), suggesting that the consistency assumption made there is actually playing a major role.

7 Conclusion

The concept of equal opportunity, with its dual principles of compensation and reward, seems to be a quite different approach to evaluating policies than the welfarist method that is standard in economics: the latter (as understood here) is based on anonymous aggregation, allowing individual characteristics to matter only insofar as they enter into the formula for individual welfare. We have asked whether this actually leads to distinguishable policy prescriptions, using a standard class of one-good distribution problems as a test case. Initial examples suggest that it does not. These examples indicate that there is no fundamental incompatibility between welfarism on the one hand, and the combination of compensation and reward on the other.

However, these examples use specific forms for the advantage function. For most possible specifications, the equivalence breaks down—the equal-opportunity choices cannot be represented by a (regular) welfarist criterion. There is a particular family of specifications for which they can; this family is characterized by the existence of AB-foliations. Moreover, within this family, the individual welfare measure is usually different from the advantage function, although in special cases it can coincide. Whenever such a welfarist representation does exist, in fact there exists a utilitarian one—where the welfare aggregator is the sum function—and also an egalitarian one—where the aggregator is the min. In the search for a welfare criterion, we have allowed ourselves complete freedom in the choice of aggregator, but this freedom turns out not to be useful.

When the sought-after welfarist representation does not exist, there is a simple structure that rules it out: a path along A-curves and B-curves from a point (c, e, x) to (c, e, x'), with $x' \neq x$. Roughly speaking, this structure rules out a utilitarian representation because, for such a representation to exist, marginal utility would have to be equal at any two points in (c, e, x)-space that can appear together in an optimal allocation; these constraints force marginal utility to be equal across too many points in the space. Similarly, an egalitarian representation is ruled out because absolute utility would have to be equal across too many points. When considering more general welfare aggregators, we can reduce to the utilitarian case by using the first-order condition for the planner's problem to show that the aggregator locally has a separable structure, and thus other aggregators do not bring us additional possibilities.

One possible interpretation of the results is that equal opportunity is indeed different from welfarism in general, but the difference is somewhat subtle. If we have particularly strong intuitions about what equal opportunity should mean in some application—in particular, if there is a path along A-curves and B-curves as above that we confidently regard as comprising the "correct" solutions to the corresponding distribution problems—then a welfarist criterion is not sufficient to capture these intuitions. On the other hand, if our intuitions are a bit less precise, it can be perfectly possible to express the ideas of compensation and reward while working in the familiar welfarist framework. In particular, the analysis offers tools for modelers, interested in writing down tractable models (such as that of Lockwood and Weinzierl (2015)) that express these dual goals; we have identified the class of advantage specifications that makes this possible. (Of course, we have worked here in the simpler setting of distribution problems; doing a similar analysis for taxation problems as in Lockwood and Weinzierl (2015), or other applications, is left to future work.)

Note also that even in the cases where a welfarist representation does exist, the individual utility measure $U(c, e, x)$ is simply a formal representation of the opportunitarian's choices in a particular class of problems. It does not follow that this measure has any natural normative interpretation; such interpretations are likely to be application-specific.

More generally, the starting point of this paper is that economics should engage more directly with the range of approaches to evaluating policies that are prominent in popular discussion, and that a first step for such engagement is to try to understand when and why these approaches have different implications than the usual welfarist approach. Identifying the differences can potentially help sharpen normative debates over which approach to use. A hope is that this paper will inspire other efforts in this direction.

A Omitted proofs

A.1 Proofs from Section 2

Proof of Proposition 2.1. First, it is a standard consequence of continuity and monotonicity that, for any X, there exists a unique quantity $x^\circ(X)$ such that the allocation that gives every agent quantity $x^\circ(X)$ is considered indifferent to X. Then, for any two allocations X, X', we have $X \succsim X'$ if and only if $x^\circ(X) \geq x^\circ(X')$, by transitivity. We use this to define a welfarist criterion as follows.

Arbitrary label the $n = |P|$ members of the population as (c_i, e_i) for $i = 1, \dots, n$. This lets us identify allocations on P with n-tuples of nonnegative numbers. For each i, let U_i be an increasing bijection from $\mathbb{R}_+$ to the interval $[i, i+1)$. Define U by $U(c_i, e_i, x) = U_i(x)$ for each i.

Let S denote the set of n-tuples of real numbers $(u_1, \dots, u_n)$ such that, for each $i = 1, \dots, n$, exactly one coordinate u_j lies in $[i, i+1)$; and for such an n-tuple and i, let $u_{(i)}$ denote the value of this coordinate. Define $\widehat{W} : S \to \mathbb{R}$ by

$$\widehat{W}(u_1, \dots, u_n) = U_1(x^\circ(U_1^{-1}(u_{(1)}), U_2^{-1}(u_{(2)}), \dots, U_n^{-1}(u_{(n)}))).$$

Because of the outer U_1 operator, $\widehat{W}$ takes values in $[1, 2)$. Now define $W : \mathbb{R}^n \to \mathbb{R}$ by

$$W(u_1, \dots, u_n) = \inf_{\substack{(u'_1, \dots, u'_n) \in S \\ u'_i \geq u_i \text{ for all } i}} \widehat{W}(u'_1, \dots, u'_n),$$

with $W(u_1, \dots, u_n) = 2$ if no such $(u'_1, \dots, u'_n)$ exists.

For completeness, we need to also define U for agents not in P and define W for arguments of dimension other than n. These can be done arbitrarily (say, $U(c, e, x) = x$ and $W(u_1, \dots, u_m) = u_1 + \cdots + u_m$).

It is immediate that U is strictly increasing in x, that W is weakly increasing (because the inf is taken over a decreasing set), and that W is symmetric. Thus $\mathcal{W} = (U, W)$ defines a welfarist criterion. We claim that for any allocation X on P, we have $\mathcal{W}(X) = U_1(x^\circ(X))$; this will imply $\mathcal{W}(X) \geq \mathcal{W}(X')$ iff $x^\circ(X) \geq x^\circ(X')$, and our desired conclusion will follow.

To verify the claim, take X given, and define $u_i = U_i(X(c_i, e_i))$ for each i. Then $(u_1, \dots, u_n) \in S$, and the only other n-tuples $(u'_1, \dots, u'_n) \in S$ that have $u'_i \geq u_i$ for each i are ones that are already in increasing order, in which case $\widehat{W}(u'_1, \dots, u'_n) \geq$

$\widehat{W}(u_1, \ldots, u_n)$. Therefore,

$$\mathcal{W}(X) = W(u_1, \ldots, u_n) = \widehat{W}(u_1, \ldots, u_n) = U_1(x^\circ(X)).$$

□

Proof of Proposition 2.2. Suppose such a criterion exists. Pick any $(c, e) \in C \times E$ and any $x \in \mathbb{R}_{++}$. Let $u^* = U(c, e, x)$. For any ϵ sufficiently small, $c + \epsilon \in C$ and $e + \epsilon \in E$, and continuity of U (and strict monotonicity in x) implies that there exist x_1, x_2 with

$$U(c + \epsilon, e, x_1) = U(c, e + \epsilon, x_2) = u^*.$$

Again by continuity and strict monotonicity in x, by choosing u' larger than u^* and sufficiently close to it, there exist x', x'_1, x'_2 with

$$U(c, e, x') = U(c + \epsilon, e, x'_1) = U(c, e + \epsilon, x'_2) = u'.$$

Consider population $\{c, c+\epsilon\} \times \{e\}$. If $A(c, e, x) \geq A(c+\epsilon, e, x_1)$, then $\mathcal{V}^{EOp}$ evaluates the two allocations $\begin{pmatrix} \overset{c,e}{x}, & \overset{c+\epsilon,e}{x_1} \end{pmatrix}$ and $\begin{pmatrix} \overset{c,e}{x'}, & \overset{c+\epsilon,e}{x_1} \end{pmatrix}$ equally (both are valued at $A(c + \epsilon, e, x_1)$). If $A(c, e, x) < A(c+\epsilon, e, x_1)$, then $\mathcal{V}^{EOp}$ evaluates $\begin{pmatrix} \overset{c,e}{x}, & \overset{c+\epsilon,e}{x_1} \end{pmatrix}$ and $\begin{pmatrix} \overset{c,e}{x}, & \overset{c+\epsilon,e}{x'_1} \end{pmatrix}$ equally (both valued at $A(c, e, x)$). So, by hypothesis, we have

$$\mathcal{W}\begin{pmatrix} \overset{c,e}{x}, & \overset{c+\epsilon,e}{x_1} \end{pmatrix} = \mathcal{W}\begin{pmatrix} \overset{c,e}{x'}, & \overset{c+\epsilon,e}{x_1} \end{pmatrix} \qquad \text{or} \qquad \mathcal{W}\begin{pmatrix} \overset{c,e}{x}, & \overset{c+\epsilon,e}{x_1} \end{pmatrix} = \mathcal{W}\begin{pmatrix} \overset{c,e}{x}, & \overset{c+\epsilon,e}{x'_1} \end{pmatrix}$$

and thus in either case

$$W(u^*, u^*) = W(u^*, u').$$

On the other hand, consider population $\{c\} \times \{e, e + \epsilon\}$. We have

$$\mathcal{V}^{EOp}\begin{pmatrix} \overset{c,e}{x}, & \overset{c,e+\epsilon}{x_2} \end{pmatrix} = A(c, e, x) + A(c, e+\epsilon, x_2) < A(c, e, x) + A(c, e+\epsilon, x'_2) = \mathcal{V}^{EOp}\begin{pmatrix} \overset{c,e}{x}, & \overset{c,e+\epsilon}{x'_2} \end{pmatrix}$$

and thus

$$\mathcal{W}\begin{pmatrix} \overset{c,e}{x}, & \overset{c,e+\epsilon}{x_2} \end{pmatrix} < \mathcal{W}\begin{pmatrix} \overset{c,e}{x}, & \overset{c,e+\epsilon}{x'_2} \end{pmatrix}$$

and thus

$$W(u^*, u^*) < W(u^*, u').$$

This is a contradiction. $\square$

A.2 Proofs from Section 3

Proof of Lemma 3.2. Consider $\epsilon > 0$ small enough that $c-\epsilon, c+\epsilon \in \overline{C}$ and $e+\epsilon \in \overline{E}$. In the distribution problem with population $\{c-\epsilon, c+\epsilon\} \times \{e\}$ and quantity $2x$, the unique equal-opportunity choice gives the two agents $x+\epsilon, x-\epsilon$; thus, for any $y \in [0, 2x]$,

$$\mathcal{W}\left(\overset{c-\epsilon,e}{x+\epsilon}, \overset{c+\epsilon,e}{x-\epsilon}\right) \geq \mathcal{W}\left(\overset{c-\epsilon,e}{2x-y}, \overset{c+\epsilon,e}{y}\right).$$

In the distribution problem with population $\{c\} \times \{e+\epsilon, e\}$ and quantity $2x$, the unique equal-opportunity choice gives all of the good to the first agent; thus,

$$\mathcal{W}\left(\overset{c,e+\epsilon}{2x}, \overset{c,e}{0}\right) \geq \mathcal{W}\left(\overset{c,e+\epsilon}{2x-y}, \overset{c,e}{y}\right).$$

Taking $\epsilon \to 0$ and using continuity gives the asserted result. $\square$

Proof of Lemma 3.3. Write $u(t) = U(\mu(t))$. By assumption, this is a continuous bijection from $[0,1]$ to $[\underline{u}, \overline{u}]$. It is either increasing or decreasing; assume it is increasing (otherwise just reverse the order on t).

Claim: If $u_1, u_1', u_2, u_2' \in [\underline{u}, \overline{u}]$ with $u_1 < u_1'$ and $u_2 < u_2'$, then $W(u_1, u_2) < W(u_1', u_2')$ strictly.

To see this, suppose otherwise. Then, because W is weakly increasing, it would have to be constant on the rectangle $[u_1, u_1'] \times [u_2, u_2']$. Then, we can choose $t \neq t'$ with $u(t) \in (u_1, u_1')$ and $u(t') \in (u_2, u_2')$. If we take any $\epsilon > 0$ sufficiently small, then still we have $U(c(t), e(t), x(t)+\epsilon) \in (u_1, u_1')$ and $U(c(t'), e(t'), x(t')-\epsilon) \in (u_2, u_2')$. But then we would have

$$\mathcal{W}\left(\overset{c(t),e(t)}{x(t)}, \overset{c(t'),e(t')}{x(t')}\right) = \mathcal{W}\left(\overset{c(t),e(t)}{x(t)+\epsilon}, \overset{c(t'),e(t')}{x(t')-\epsilon}\right),$$

contrary to condition 2 of the lemma.

Now, define H by

$$H(u) = \int_{\underline{u}}^{u} \frac{1}{\frac{\partial U}{\partial x}\Big|_{\mu(u^{-1}(\widetilde{u}))}} \, d\widetilde{u}.$$

Thus, H is a differentiable, strictly increasing bijection from $[\underline{u}, \overline{u}]$ to $[0, m]$, where $m =$

$H(\overline{u})$, and we have

$$H'(u(t)) \times \left.\frac{\partial U}{\partial x}\right|_{\mu(t)} = 1$$

for each $t \in [0,1]$.

We wish to show that $W(u_1, u_2)$ is a function of $H(u_1) + H(u_2)$. Suppose for contradiction that there exist two points $(u_1, u_2), (u_1', u_2') \in (\underline{u}, \overline{u})^2$ such that

$$H(u_1) + H(u_2) = H(u_1') + H(u_2') \tag{A.1}$$

but $W(u_1, u_2) \neq W(u_1', u_2')$. Without loss of generality, assume $W(u_1', u_2') < W(u_1, u_2)$. In order for (A.1) to hold, it must be that $u_1' < u_1$ and $u_2' > u_2$ or vice versa. Assume that indeed $u_1' < u_1$ and $u_2' > u_2$ (otherwise, just swap u_1 with u_2 and u_1' with u_2'). Now consider the assertion

$$W\left(H^{-1}(H(u_1) + H(u_2) - (1-\delta)H(u_2')), u_2'\right) < W(u_1, u_2). \tag{A.2}$$

When $\delta = 0$, the left side is simply $W(u_1', u_2')$, and the assertion holds by hypothesis. So, for small $\delta > 0$, the left side remains well-defined and (A.2) continues to hold. Fix such a δ.

Now define the differentiable function J on the rectangle $[\underline{u}, \overline{u}]^2$ by

$$J(u_1'', u_2'') = H(u_1'') + (1-\delta)H(u_2'').$$

Choose a point (u_1'', u_2'') to maximize the value of J over the rectangle $[u_1', u_1] \times [u_2, u_2']$, subject to the constraint $W(u_1'', u_2'') \leq W(u_1, u_2)$. From (A.2), the point $(H^{-1}(H(u_1) + H(u_2) - (1-\delta)H(u_2')), u_2')$ is feasible in this maximization problem, and the value of J there is higher than $J(u_1, u_2)$, implying that the point (u_1, u_2) is not optimal in the maximization problem. Moreover, J is increasing in each coordinate. Consequently, we must have $u_2'' > u_2$.

We have either $u_1'' < u_1$ or $u_1'' = u_1$. Suppose first that $u_1'' < u_1$. Consider points t_1 and t_2 such that $u(t_1) = u_1''$ and $u(t_2) = u_2''$. For small $\epsilon > 0$, define

$$\widetilde{u}_1(\epsilon) = U(c(t_1), e(t_1), x(t_1) + \epsilon);$$

$$\widetilde{u}_2(\epsilon) = U(c(t_2), e(t_2), x(t_2) - \epsilon).$$

For small ϵ, the point $(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon))$ still lies in the rectangle $[u_1', u_1] \times [u_2, u_2']$, and

$$W(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon)) = \mathcal{W}\begin{pmatrix} \overset{c(t_1),e(t_1)}{x(t_1)+\epsilon}, \overset{c(t_2),e(t_2)}{x(t_2)-\epsilon} \end{pmatrix} \leq \mathcal{W}\begin{pmatrix} \overset{c(t_1),e(t_1)}{x(t_1)}\ ,\ \overset{c(t_2),e(t_2)}{x(t_2)} \end{pmatrix} = W(u_1'', u_2''), \tag{A.3}$$

where the inequality again comes from hypothesis 2 of the lemma. So $(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon))$ is a feasible point in the maximization problem defining (u_1'', u_2''). Moreover,

$$\begin{aligned} \frac{d}{d\epsilon} J(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon))\bigg|_{\epsilon=0} &= H'(u(t_1)) \times \frac{\partial U}{\partial x}\bigg|_{\mu(t_1)} + (1-\delta) \times H'(u(t_2)) \times \left(-\frac{\partial U}{\partial x}\bigg|_{\mu(t_2)}\right) \\ &= 1 + (1-\delta) \times (-1) \\ &= \delta \\ &> 0. \end{aligned}$$

Therefore, for small ϵ, we have $J(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon)) > J(\widetilde{u}_1(0), \widetilde{u}_2(0)) = J(u_1'', u_2'')$. This contradicts the optimality in the definition of (u_1'', u_2'').

Therefore, we must instead have $u_1'' = u_1$. In this case, again define $t_1, t_2, \widetilde{u}_1(\epsilon)$, and $\widetilde{u}_2(\epsilon)$ as in the previous case. For small $\epsilon > 0$, we still have inequality (A.3). Now, however, the point $(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon))$ lies strictly above (u_1, u_2) in both coordinates. However, from our initial claim, this implies $W(u_1, u_2) < W(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon))$. Thus,

$$W(u_1, u_2) < W(\widetilde{u}_1(\epsilon), \widetilde{u}_2(\epsilon)) \leq W(u_1'', u_2'') \leq W(u_1, u_2),$$

which is impossible. So this case leads to a contradiction too.

Thus, it follows that for all $(u_1, u_2) \in (\underline{u}, \overline{u})^2$, the value of $H(u_1) + H(u_2)$ uniquely determines $W(u_1, u_2)$, i.e. we can write

$$W(u_1, u_2) = G(H(u_1) + H(u_2)) \tag{A.4}$$

for some function G. As u_1 varies continuously over the interval $(\underline{u}, \overline{u})$, and u_2 is kept equal to u_1, the value of $H(u_1) + H(u_2)$ increases continuously from 0 to $2m$, and $W(u_1, u_2)$ also increases continuously (strictly, by the claim); this pins down G throughout the interval $(0, 2m)$, and G is continuous and strictly increasing. Finally, we can extend the definition of G to the endpoints by $G(0) = W(\underline{u}, \underline{u})$ and $G(2m) = W(\overline{u}, \overline{u})$; easily, (A.4) implies G is continuous at these endpoints. Then, by continuity (A.4) holds on the boundary of the square $[\underline{u}, \overline{u}]$ as well, so it holds throughout the square. And we have already checked that G and H are both strictly increasing and H is differentiable, as needed. □

Proof of Theorem 3.1. We assume the desired welfarist criterion $\mathcal{W} = (U, W)$ exists, and seek a contradiction.

First we show that one of the following two cases holds:

(i) There exist $c^* \in \overline{C}$ and $e^* \in \overline{E}$, and small $\epsilon > 0$, such that $U(c, e^*, 0)$ is constant in c for $c \in [c^* - \epsilon, c^* + \epsilon]$.

(ii) There exist $c^* \in \overline{C}$, $e^* \in \overline{E}$, $\epsilon > 0$, and values $\underline{u} < \overline{u}$ such that

- $W|_{\mathbb{R}^2}$ has an additive representation (G, H) on the interval $[\underline{u}, \overline{u}]$, with G, H strictly increasing, and
- $U(c, e, x) \in (\underline{u}, \overline{u})$ for all $(c, e, x) \in [c^* - \epsilon, c^* + \epsilon] \times \{e^*\} \times [0, 4\epsilon]$.

(In both cases, it is understood that ϵ is small enough so that $[c^* - \epsilon, c^* + \epsilon] \subseteq \overline{C}$.)

To see this, suppose (i) does not hold. Fix any $e^* \in \overline{E}$. Choose $c^\dagger$ such that $\partial U/\partial c|_{(c^\dagger, e^*, 0)} \neq 0$. (We can do this, since U is differentiable and, by assumption, non-constant in c for $x = 0$.) Choose $\epsilon > 0$ small enough so that $c^\dagger - 2\epsilon > \underline{c}$ and $\partial U/\partial c|_{(c^\dagger - \epsilon, e^*, 0)} \neq 0$.

We claim that one of the choices $c^* = c^\dagger$ or $c^* = c^\dagger - \epsilon$ has the following property: the function $U(c, e^*, c^* - c)$ is non-constant in c on the interval $(\underline{c}, c^*]$. Suppose this is not true for either choice. The constancy assumption implies

$$U(c^\dagger - 2\epsilon, e^*, \epsilon) = U(c^\dagger - \epsilon, e^*, 0) \tag{A.5}$$

and

$$U(c^\dagger - 2\epsilon, e^*, 2\epsilon) = U(c^\dagger - \epsilon, e^*, \epsilon), \tag{A.6}$$

and therefore

$$\mathcal{W}\begin{pmatrix} c^\dagger-2\epsilon, e^* & c^\dagger-\epsilon, e^* \\ 2\epsilon \;, & \epsilon \end{pmatrix} = \mathcal{W}\begin{pmatrix} c^\dagger-\epsilon, e^* & c^\dagger-\epsilon, e^* \\ \epsilon \;, & \epsilon \end{pmatrix} = \mathcal{W}\begin{pmatrix} c^\dagger-\epsilon, e^* & c^\dagger-\epsilon, e^* \\ 0 \;, & 2\epsilon \end{pmatrix} = \mathcal{W}\begin{pmatrix} c^\dagger-2\epsilon, e^* & c^\dagger-\epsilon, e^* \\ \epsilon \;, & 2\epsilon \end{pmatrix}.$$

Here, the first equality is by (A.6), the second by Lemma 3.2, and the third by (A.5). However, in the distribution problem with population $\{c^\dagger - 2\epsilon, c^\dagger - \epsilon\} \times \{e^*\}$ and quantity 3ϵ, the allocation that gives the agents $2\epsilon, \epsilon$ respectively is an equal-opportunity choice, while the allocation that gives them $\epsilon, 2\epsilon$ is not, so $\mathcal{W}$ should not value these allocations equally—a contradiction.

This proves the claim. Thus we now have c^*, e^* such that $\partial U/\partial c|_{(c^*,e^*,0)} \neq 0$, and $U(c, e^*, c^* - c)$ is non-constant for $c \in (\underline{c}, c^*]$. The former property implies that, taking $u^* = U(c^*, e^*, 0)$, for small enough $\delta > 0$ we have

$$(u^* - \delta, u^* + \delta) \subseteq \{U(c', e^*, 0) \mid c' \in \overline{C}\}.$$

Using the latter property, let $R(c) = U(c, e^*, c^* - c)$; because R is differentiable and non-constant in c, we can choose some point where its derivative is nonzero, and thus have a closed interval $I \subseteq (\underline{c}, c^*)$ mapped bijectively by R to an interval $[\underline{u}, \overline{u}] \subseteq \mathbb{R}$. Moreover, we can choose this interval so that its image under R reaches points arbitrarily close to $U(c^*, e^*, 0) = u^*$, in particular, so that

$$(\underline{u}, \overline{u}) \cap (u^* - \delta, u^* + \delta) \neq \emptyset. \tag{A.7}$$

For any distinct $c < c' \in I$, we can consider the distribution problem with population $\{c, c'\} \times \{e^*\}$ and quantity $2c^* - (c + c')$. The unique equal-opportunity allocation gives these agents $c^* - c$ and $c^* - c'$. Thus, for nonzero ϵ with $|\epsilon| < c^* - c'$, we have

$$\mathcal{W}\left(\overset{c,e^*}{c^* - c},\ \overset{c',e^*}{c^* - c'}\right) > \mathcal{W}\left(\overset{c,e^*}{c^* - c + \epsilon},\ \overset{c',e^*}{c^* - c' - \epsilon}\right).$$

The same inequality holds weakly for $c' = c$, by continuity (or by Lemma 3.2). Thus, all the conditions to apply Lemma 3.3 are satisfied, and we obtain the claimed additive representation.

This covers the first statement of (ii). For the second, by (A.7), we can find a point c' with $U(c', e^*, 0) \in (\underline{u}, \overline{u})$. By continuity, U remains in $(\underline{u}, \overline{u})$ throughout a neighborhood of $(c', e^*, 0)$. Thus, relabeling c' as c^* (which does not affect the first statement), we get the second statement of (ii) as well.

Thus we have shown that one of the cases (i), (ii) must hold, and now we proceed to obtain a contradiction in each case.

First suppose case (i) holds. Then,

$$\mathcal{W}\left(\overset{c^*,e^*}{\epsilon},\ \overset{c^*+\epsilon,e^*}{0}\right) > \mathcal{W}\left(\overset{c^*,e^*}{0},\ \overset{c^*+\epsilon,e^*}{\epsilon}\right) = \mathcal{W}\left(\overset{c^*+\epsilon,e^*}{0},\ \overset{c^*+\epsilon,e^*}{\epsilon}\right) = \mathcal{W}\left(\overset{c^*+\epsilon,e^*}{\epsilon},\ \overset{c^*+\epsilon,e^*}{0}\right).$$

The inequality is by representation of equal opportunity, the first equality by the assumption of case (i), and the second by symmetry.

By monotonicity of W, this implies

$$U(c^*, e^*, \epsilon) > U(c^* + \epsilon, e^*, \epsilon). \tag{A.8}$$

We also have

$$\mathcal{W}\left(\overset{c^*,e^*}{2\epsilon}, \overset{c^*+\epsilon,e^*}{\epsilon}\right) > \mathcal{W}\left(\overset{c^*,e^*}{3\epsilon}, \overset{c^*+\epsilon,e^*}{0}\right) = \mathcal{W}\left(\overset{c^*,e^*}{3\epsilon}, \overset{c^*,e^*}{0}\right) \geq \mathcal{W}\left(\overset{c^*,e^*}{2\epsilon}, \overset{c^*,e^*}{\epsilon}\right)$$

where the first inequality is by representation of equal opportunity, the equality by case (i), and the second inequality by Lemma 3.2. Then by monotonicity of W we get

$$U(c^* + \epsilon, e^*, \epsilon) > U(c^*, e^*, \epsilon)$$

contradicting (A.8).

Now suppose case (ii) holds. Let (G, H) denote the additive representation. We hold fixed $e = e^*$ throughout the rest of the proof. For any $c \in [c^* - \epsilon, c^* + \epsilon]$ and $x, y \in [0, 4\epsilon]$, it will be convenient to write

$$\Delta\left(x \overset{c}{\to} y\right) = H(U(c, e^*, y)) - H(U(c, e^*, x)).$$

To see the usefulness of this notation, note that whenever

$$\mathcal{W}\left(\overset{c,e^*}{x}, \overset{c',e^*}{x'}\right) \geq \mathcal{W}\left(\overset{c,e^*}{y}, \overset{c',e^*}{y'}\right), \tag{A.9}$$

we can expand using the additive representation of W, cancel G's on both sides, and then rearrange terms to get

$$\Delta\left(y' \overset{c'}{\to} x'\right) \geq \Delta\left(x \overset{c}{\to} y\right). \tag{A.10}$$

Likewise, if (A.9) holds as an equality (resp. strict equality), so does (A.10).

From Lemma 3.2, we have

$$\mathcal{W}\left(\overset{c^*-\epsilon,e^*}{\epsilon}, \overset{c^*-\epsilon,e^*}{\epsilon}\right) = \mathcal{W}\left(\overset{c^*-\epsilon,e^*}{0}, \overset{c^*-\epsilon,e^*}{2\epsilon}\right),$$

and applying the above observation gives us

$$\Delta\left(\epsilon \overset{c^*-\epsilon}{\rightarrow} 2\epsilon\right) = \Delta\left(0 \overset{c^*-\epsilon}{\rightarrow} \epsilon\right).$$

Similar reasoning, again using Lemma 3.2, gives

$$\Delta\left(2\epsilon \overset{c^*-\epsilon}{\rightarrow} 4\epsilon\right) = \Delta\left(0 \overset{c^*-\epsilon}{\rightarrow} 2\epsilon\right)$$

and

$$\Delta\left(\epsilon \overset{c^*}{\rightarrow} 2\epsilon\right) = \Delta\left(0 \overset{c^*}{\rightarrow} \epsilon\right).$$

Now, in the distribution problem with population $\{c^* - \epsilon, c^*\} \times \{e^*\}$ and quantity ϵ, representation of equal opportunity implies

$$\mathcal{W}\left(\overset{c^*-\epsilon, e^*}{\epsilon}, \overset{c^*, e^*}{0}\right) > \mathcal{W}\left(\overset{c^*-\epsilon, e^*}{0}, \overset{c^*, e^*}{\epsilon}\right),$$

thus

$$\Delta\left(0 \overset{c^*-\epsilon}{\rightarrow} \epsilon\right) > \Delta\left(0 \overset{c^*}{\rightarrow} \epsilon\right). \tag{A.11}$$

By similar analysis for the problem with same population and quantity 3ϵ, we get

$$\Delta\left(0 \overset{c^*}{\rightarrow} \epsilon\right) > \Delta\left(2\epsilon \overset{c^*-\epsilon}{\rightarrow} 3\epsilon\right)$$

and, for the same population with quantity 5ϵ, we get

$$\Delta\left(\epsilon \overset{c^*}{\rightarrow} 2\epsilon\right) > \Delta\left(3\epsilon \overset{c^*-\epsilon}{\rightarrow} 4\epsilon\right).$$

Combining,

$$\begin{aligned}
\Delta\left(0 \overset{c^*}{\rightarrow} 2\epsilon\right) &= \Delta\left(0 \overset{c^*}{\rightarrow} \epsilon\right) + \Delta\left(\epsilon \overset{c^*}{\rightarrow} 2\epsilon\right) \\
&> \Delta\left(2\epsilon \overset{c^*-\epsilon}{\rightarrow} 3\epsilon\right) + \Delta\left(3\epsilon \overset{c^*-\epsilon}{\rightarrow} 4\epsilon\right) \\
&= \Delta\left(2\epsilon \overset{c^*-\epsilon}{\rightarrow} 4\epsilon\right) \\
&= \Delta\left(0 \overset{c^*-\epsilon}{\rightarrow} 2\epsilon\right).
\end{aligned}$$

However,

$$\Delta\left(0 \xrightarrow{c^*} 2\epsilon\right) = \Delta\left(0 \xrightarrow{c^*} \epsilon\right) + \Delta\left(\epsilon \xrightarrow{c^*} 2\epsilon\right) = 2\Delta\left(0 \xrightarrow{c^*} \epsilon\right)$$

and likewise

$$\Delta\left(0 \xrightarrow{c^*-\epsilon} 2\epsilon\right) = 2\Delta\left(0 \xrightarrow{c^*-\epsilon} \epsilon\right),$$

so we conclude

$$\Delta\left(0 \xrightarrow{c^*} \epsilon\right) > \Delta\left(0 \xrightarrow{c^*-\epsilon} \epsilon\right),$$

contradicting (A.11). □

A.3 Proofs from Section 4

Proof of Lemma 4.3. Part 1: Continuity ensures a maximizer exists; we show any maximizer X must have the stated property. Let $A^* = \min_c A(c, e, X(c, e))$. If c' is such that $A(c', e, X(c', e)) > A^*$ and $X(c', e) > 0$, then we can decrease the quantity given to agent (c', e) by a small amount ϵ and increase the quantity given to each other agent by $\frac{\epsilon}{|C|-1}$, thereby increasing the objective, contradicting optimality. So every agent either is given enough to exactly reach advantage level A^*, or is given zero and already has $A(c', e, 0) > A^*$, which implies the stated property.

We also need to show that there is only one allocation with this property. Suppose that two different allocations X^*, X^{**} have the stated property, with two different advantage values $A^* < A^{**}$. Then, in particular, every agent with $A(c, e, 0) < A^*$ receives positive quantities in both allocations, and strictly more in allocation X^{**} than X^*. But in X^*, these agents' quantities already add up to the total $\overline{x}$ (since the remaining agents receive 0). So their quantities in X^{**} add up to more than $\overline{x}$, which is impossible.

Part 2: Again, a maximizer exists, and we show any maximizer must have the stated property. Let $B^* = \min_e B(c, e, X(c, e))$, and pick some e attaining the min. If e' is such that $B(c, e', X(c, e')) > B^*$ and $X(c, e') > 0$, then we can change the quantities assigned to agents (c, e) and (c, e') by $+\epsilon$ and $-\epsilon$, and increase the objective, a contradiction. Thus, every agent either is given enough to reach a B-value of B^*, or is given zero and already has $B(c, e', 0) > B^*$, implying the stated property.

The proof that only one allocation has the stated property is exactly the same as for part 1. □

For the remainder of this appendix section, we maintain Assumption 4.1.

For the proofs of Theorems 4.4 and 4.5, it will be useful to have a few more definitions.

For $(c,e,x),(c',e',x') \in \overline{C}\times\overline{E}\times\mathbb{R}_{++}$, write $(c,e,x) \leftrightarrow (c',e',x')$ if the points $(c,e,x),(c',e',x')$ lie on the same A-curve or the same B-curve.

Suppose $\overline{C}' \subseteq \overline{C}$, $\overline{E}' \subseteq \overline{E}$ are open intervals, and $\xi : \overline{C}' \times \overline{E}' \to \mathbb{R}_{++}$ is a function. Say that ξ *locally respects AB-foliation* if, for all $(c,e),(c',e') \in \overline{C}' \times \overline{E}'$, if $c = c'$ or $e = e'$ then $(c,e,\xi(c,e)) \leftrightarrow (c',e',\xi(c',e'))$. A *local foliation leaf* is any set of the form

$$\{(c,e,\xi(c,e)) \mid (c,e) \in \overline{C}' \times \overline{E}'\},$$

where $\xi : \overline{C}' \times \overline{E}' \to \mathbb{R}_{++}$ locally respects AB-foliation.

If (U,W) is a regular welfarist criterion, a *local level set* of U is a set of the form

$$\{(c,e,x) \mid (c,e) \in \overline{C}' \times \overline{E}',\ U(c,e,x) = u^*\}$$

where $\overline{C}', \overline{E}'$ are open intervals and u^* is a value such that, for every $(c,e) \in \overline{C}' \times \overline{E}'$, some $x \in \mathbb{R}_{++}$ with $U(c,e,x) = u^*$ exists. Note that for every $(c^*,e^*,x^*) \in \overline{C} \times \overline{E} \times \mathbb{R}_{++}$, there exists a local level set of U containing it, by choosing $\overline{C}', \overline{E}'$ sufficiently small.

For any $(c,e,x) \in \overline{C} \times \overline{E} \times \mathbb{R}_{++}$, and any $c' \in \overline{C}$, define $\xi^A(c'|c,e,x)$ to be the value such that

$$A(c',e,\xi^A(c'|c,e,x)) = A(c,e,x)$$

if such a value exists. Note that ξ^A is defined on a neighborhood of $\{(c',c,e,x) \mid c' = c\}$, and by the implicit function theorem, it is differentiable there, with

$$\frac{\partial \xi^A}{\partial c'} = -\left.\frac{\partial A/\partial c}{\partial A/\partial x}\right|_{(c',e,\xi^A(c'|c,e,x))}.$$

Similarly, for $e' \in \overline{E}$, define $\xi^B(e'|c,e,x)$ to be the value such that $B(c,e',\xi^B(e'|c,e,x)) = B(c,e,x)$ if such a value exists; ξ^B is defined on a neighborhood of $\{(e',c,e,x) \mid e' = e\}$, and differentiable with

$$\frac{\partial \xi^B}{\partial e'} = -\left.\frac{\partial B/\partial e}{\partial B/\partial x}\right|_{(c,e',\xi^B(e'|c,e,x))}.$$

Notice also that if the function $\xi^A(\cdot|c,e,x)$ is defined at some point c' and at some other point c'', then it is also defined for every value in between, by the monotonicity of A with respect to c in Assumption 4.1 (and continuity in x). Likewise for ξ^B.

Lemma A.1. *Suppose the regular welfarist criterion $\mathcal{W} = (U,W)$ represents equal opportunity for two agents. Suppose that S is a local level set of U, with U-value u^*. If S is not* *a local foliation leaf, then $W|_{\mathbb{R}^2}$ has an additive representation, satisfying the conclusions of Lemma 3.3, on some interval $[\underline{u}, \overline{u}]$ with $\underline{u} < u^* < \overline{u}$.*

Proof. Let ξ be the function such that $x = \xi(c, e)$ for each $(c, e, x) \in S$. By the implicit function theorem, ξ is differentiable, with $\partial\xi/\partial c = -\frac{\partial U/\partial c}{\partial U/\partial x}$ and $\partial\xi/\partial e = -\frac{\partial U/\partial e}{\partial U/\partial x}$.

Because S is not a local foliation leaf, there exist $(c', e', x'), (c'', e'', x'') \in S$ with $c' = c''$ or $e' = e''$ but $(c', e', x') \not\leftrightarrow (c'', e'', x'')$. Assume $e' = e''$; the $c' = c''$ case is analogous. Fix $e^\circ = e'$.

Thus, $A(c', e^\circ, x') \neq A(c'', e^\circ, x'')$. Consquently, there exists a point c° where the (total) derivative of $A(c, e^\circ, \xi(c, e^\circ))$ with respect to c is nonzero:

$$\frac{\partial A}{\partial c} + \frac{\partial A}{\partial x} \cdot \left(-\frac{\partial U/\partial c}{\partial U/\partial x}\right) \neq 0 \qquad \text{at } (c^\circ, e^\circ, \xi(c^\circ, e^\circ)).$$

Put $x^\circ = \xi(c^\circ, e^\circ)$. By rearranging, the total derivative of the function $U^\dagger(c) = U(c, e^\circ, \xi^A(c|c^\circ, e^\circ, x^\circ))$ with respect to c is nonzero at c°:

$$\frac{\partial U}{\partial c} + \frac{\partial U}{\partial x} \cdot \left(-\frac{\partial A/\partial c}{\partial A/\partial x}\right) \neq 0 \qquad \text{at } (c^\circ, e^\circ, x^\circ).$$

Then, there exists a closed interval of c-values with c° in its interior, such that the function $U^\dagger$ takes this interval bijectively to an interval $[\underline{u}, \overline{u}]$ in $\mathbb{R}$. Letting $t \in [0, 1]$ parameterize the interval, define

$$\mu(t) = (c(t), e(t), x(t)) = (c(t), e^\circ, \xi^A(c(t)|c^\circ, e^\circ, x^\circ)).$$

So condition 1 of Lemma 3.3 is met. To check condition 2, note that for any $t \neq t'$, we can take the distribution problem with population $P = \{c(t), c(t')\} \times \{e^\circ\}$, and total quantity $\overline{x} = x(t) + x(t')$. By Lemma 4.3, the unique equal-opportunity choice in this problem gives the agents quantities $x(t), x(t')$ respectively (since they then attain equal values of A). Since $\mathcal{W}$ represents equal opportunity for two agents, it follows that the strict inequality in condition 2 is met, with $\overline{\epsilon} = \min_t x(t)$ (so that the allocations in the condition are defined). Finally, continuity ensures that condition 2 is also satisfied (as a weak inequality) when $t' = t$.

Thus Lemma 3.3 applies, giving the desired additive representation on $[\underline{u}, \overline{u}]$. And $u^* = U(c^\circ, e^\circ, x^\circ) = U^\dagger(c^\circ)$ is in the interior of this interval. □

Lemma A.2. *Suppose that the regular welfarist criterion $\mathcal{W} = (U, W)$ represents equal opportunity for two agents. Suppose that, on some interval $[\underline{u}, \overline{u}]$, $W|_{\mathbb{R}^2}$ has an additive* *representation (G, H) satisfying the conclusion of Lemma 3.3. Suppose that (c_1, e_1, x_1) and (c_2, e_2, x_2) are two points on the same A-curve or B-curve, with $x_1, x_2 > 0$, and suppose that $u_1 = U(c_1, e_1, x_1)$ and $u_2 = U(c_2, e_2, x_2)$ both lie in the interval $(\underline{u}, \overline{u})$. Then,*

$$H'(u_1) \times \left.\frac{\partial U}{\partial x}\right|_{(c_1,e_1,x_1)} = H'(u_2) \times \left.\frac{\partial U}{\partial x}\right|_{(c_2,e_2,x_2)}.$$

Proof. The assumptions imply that $P = \{(c_1, e_1), (c_2, e_2)\}$ is a population, and in the distribution problem with population P and total quantity $x_1 + x_2$, the unique equal-opportunity choice gives the agents x_1 and x_2. So, if $\mathcal{W}$ represents equal opportunity for two agents, then

$$\mathcal{W}\left(\overset{c_1,e_1}{x_1},\ \overset{c_2,e_2}{x_2}\right) \geq \mathcal{W}\left(\overset{c_1,e_1}{x_1+\epsilon},\ \overset{c_2,e_2}{x_2-\epsilon}\right)$$

for all $|\epsilon| < \min\{x_1, x_2\}$. In particular, when ϵ is close enough to 0, $U(c_1, e_1, x_1 + \epsilon), U(c_2, e_2, x_2 - \epsilon) \in (\underline{u}, \overline{u})$, and we can use the additive representation, then apply G^{-1} to get

$$H(U(c_1, e_1, x_1)) + H(U(c_2, e_2, x_2)) \geq H(U(c_1, e_1, x_1 + \epsilon)) + H(U(c_2, e_2, x_2 - \epsilon)). \quad \text{(A.12)}$$

That is, the right side of (A.12) is locally maximized at $\epsilon = 0$. The first-order condition then gives the desired result. □

Lemma A.3. *Suppose that the regular welfarist criterion $\mathcal{W} = (U, W)$ represents equal opportunity for two agents. Then, every point $(c^*, e^*, x^*) \in \overline{C} \times \overline{E} \times \mathbb{R}_{++}$ is contained in some local foliation leaf.*

Proof. Fix (c^*, e^*, x^*), and let $u^* = U(c^*, e^*, x^*)$. Let S be a local level set of U around (c^*, e^*, x^*). If S is a local foliation leaf then we are done. Suppose not; then Lemma A.1 applies. Let $[\underline{u}, \overline{u}]$ be as given by the lemma, and let (G, H) be the corresponding additive representation for W there.

Consider neighborhoods $\overline{C}'$ of c^* in $\overline{C}$, and $\overline{E}'$ of e^* in $\overline{E}$. For $(c, e) \in \overline{C}' \times \overline{E}'$, consider attempting to define

$$\begin{aligned} x_1(c, e) &= \xi^A(c|c^*, e^*, x^*); \\ x_2(c, e) &= \xi^B(e|c, e^*, x_1(c, e)); \\ x_3(c, e) &= \xi^A(c^*|c, e, x_2(c, e)); \\ x_4(c, e) &= \xi^B(e^*|c^*, e, x_3(c, e)). \end{aligned}$$

These may not be defined for all (c, e). However, by straightforward induction on j, each function x_j is defined and continuous on some neighborhood of (c^*, e^*), and takes value x^* at (c^*, e^*). So, by taking $\overline{C}' \times \overline{E}'$ to be a sufficiently small neighborhood of (c^*, e^*), we can assume that all x_j are defined and continuous there. Moreover, if $\overline{C}', \overline{E}'$ are sufficiently small, and $\overline{X}' \subseteq \mathbb{R}_{++}$ is a sufficiently small neighborhood of x^*, then the image of U on $\overline{C}' \times \overline{E}' \times \overline{X}'$ is contained in $[\underline{u}, \overline{u}]$. Assume this is the case; and by shrinking $\overline{C}', \overline{E}'$ again if necessary, we can further assume that for $(c, e) \in \overline{C}' \times \overline{E}'$, all values of $x_j(c, e)$ are contained in $\overline{X}'$.

There are two possibilities: either

(i) for all $(c, e) \in \overline{C}' \times \overline{E}'$, we have $x_4(c, e) = x^*$; or

(ii) there is some $(c, e) \in \overline{C}' \times \overline{E}'$ for which this is not the case.

In case (i), we claim that x_2 locally respects AB-foliation (on $\overline{C}' \times \overline{E}'$); this then gives us our local foliation leaf. To check this, consider two points $(c, e), (c', e') \in \overline{C}' \times \overline{E}'$. If $c = c'$, then $x_1(c, e) = x_1(c, e')$, so

$$B(c, e, x_2(c, e)) = B(c, e^*, x_1(c, e)) = B(c, e^*, x_1(c, e')) = B(c, e', x_2(c, e'))$$

and hence $(c, e, x_2(c, e)) \leftrightarrow (c, e', x_2(c, e'))$. If $e = e'$, then using (i), $x_4(c, e) = x^* = x_4(c', e)$ implies

$$x_3(c, e) = \xi^B(e|c^*, e^*, x_4(c, e)) = \xi^B(e|c^*, e^*, x_4(c', e)) = x_3(c', e)$$

and so

$$A(c, e, x_2(c, e)) = A(c^*, e, x_3(c, e)) = A(c^*, e, x_3(c', e)) = A(c', e, x_2(c', e)),$$

hence $(c, e, x_2(c, e)) \leftrightarrow (c', e, x_2(c', e))$.

This leaves case (ii), and we will show that this case leads to a contradiction. Define K to be the partial derivative of $H \circ U$ with respect to x, on the domain $\overline{C}' \times \overline{E}' \times \overline{X}'$; thus

$$K(c, e, x) = H'(U(c, e, x)) \cdot \left.\frac{\partial U}{\partial x}\right|_{(c,e,x)}$$

and this definition makes sense on its domain. Lemma A.2 tells us that, whenever two points $(c, e, x), (c', e', x') \in \overline{C}' \times \overline{E}' \times \overline{X}'$ lie on the same A-curve or the same B-curve, then $K(c, e, x) = K(c', e', x')$. In particular, for any $(c, e) \in \overline{C}' \times \overline{E}'$, this implies

$$K(c^*, e^*, x^*) = K(c, e^*, x_1(c, e)) = K(c, e, x_2(c, e)) = K(c^*, e, x_3(c, e)) = K(c^*, e^*, x_4(c, e)). \tag{A.13}$$

Now, by hypothesis of case (ii), there is (c, e) such that $x_4(c, e) \neq x^*$; put $x^{**} = x_4(c, e)$. So we have

$$K(c^*, e^*, x^*) = K(c^*, e^*, x^{**});$$

call this shared value K^*. Moreover, for each $t \in [0, 1]$, repeating (A.13) with $(tc + (1 - t)c^*, te + (1 - t)e^*)$ in place of (c, e) gives

$$K(c^*, e^*, x_4(tc + (1 - t)c^*, te + (1 - t)e^*)) = K^*$$

for each t. Now, $x_4(tc + (1 - t)c^*, te + (1 - t)e^*)$ is continuous in t, so as t varies between 0 and 1, it ranges over (at least) the interval between x^* and x^{**}; thus for each x in this interval we have

$$K(c^*, e^*, x) = K^*.$$

By choosing $\delta > 0$ small, we also have $\xi^A(c^* + \delta | c^*, e^*, x^*)$ and $\xi^A(c^* + \delta | c^*, e^*, x^{**})$ lying in $\overline{X}'$, and distinct. Fix such a δ. Put $\zeta(x) = \xi^A(c^* + \delta | c^*, e^*, x)$, which is increasing in x. Again by continuity, as x ranges between x^* and x^{**}, $\zeta(x)$ ranges over the interval between $\zeta(x^*)$ and $\zeta(x^{**})$. Moreover, for each x, Lemma A.2 again applies to tell us

$$K(c^* + \delta, e^*, \zeta(x)) = K(c^*, e^*, x) = K^*.$$

Fix $x^\dagger = (x^* + x^{**})/2$. For small enough $\bar{\epsilon}$, the above imply

$$K(c^*, e^*, x^\dagger + \epsilon) = K(c^* + \delta, e^*, \zeta(x^\dagger) - \epsilon) = K^* \qquad \text{for all } \epsilon \in (-\bar{\epsilon}, \bar{\epsilon}).$$

Therefore, if we consider the various allocations

$$\left(\overset{c^*, e^*}{x^\dagger + \epsilon}, \overset{c^* + \delta, e^*}{\zeta(x^\dagger) - \epsilon} \right)$$

for $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon})$, all of them consist of points where the additive representation of W applies and where $K = K^*$. Therefore, for each such ϵ, we have

$$\mathcal{W}\left(\overset{c^*,e^*}{x^\dagger + \epsilon}, \overset{c^*+\delta,e^*}{\zeta(x^\dagger) - \epsilon}\right) = G(H(U(c^*, e^*, x^\dagger + \epsilon)) + H(U(c^* + \delta, e^*, \zeta(x^\dagger) - \epsilon))). \quad \text{(A.14)}$$

Note that

$$\frac{d}{d\epsilon}\left(H(U(c^*, e^*, x^\dagger + \epsilon)) + H(U(c^* + \delta, e^*, \zeta(x^\dagger) - \epsilon))\right)$$
$$= K(c^*, e^*, x^\dagger + \epsilon) - K(c^* + \delta, e^*, \zeta(x^\dagger) - \epsilon) = K^* - K^* = 0$$

and therefore (A.14) is constant in $\epsilon \in (-\overline{\epsilon}, \overline{\epsilon})$.

However, if we consider the distribution problem with population $\{c^*, c^* + \delta\} \times \{e^*\}$ and total quantity $x^\dagger + \zeta(x^\dagger)$, the unique equal-opportunity choice gives the two agents quantities $x^\dagger$ and $\zeta(x^\dagger)$. Since $\mathcal{W}$ represents equal opportunity for two agents, this implies that the value of (A.14) should be strictly higher at $\epsilon = 0$ than any other nearby ϵ. But we just showed that (A.14) is locally constant in ϵ—a contradiction. □

Lemma A.4. *Suppose that every point in $\overline{C} \times \overline{E} \times \mathbb{R}_{++}$ is contained in some local foliation leaf. Then, an AB-foliation exists.*

Proof. Define the binary relation $\approx$, on $\overline{C} \times \overline{E} \times \mathbb{R}_{++}$, as the transitive closure of the relation $\leftrightarrow$. It is clear that $\approx$ is an equivalence relation, and that if $(c, e, x), (c', e', x')$ are two points with $(c, e, x) \leftrightarrow (c', e', x')$ then $(c, e, x) \approx (c', e', x')$. We need to show the converse: if $(c, e, x) \approx (c', e', x')$, then $e = e'$ implies $A(c, e, x) = A(c', e', x')$, and $c = c'$ implies $B(c, e, x) = B(c', e', x')$.

So suppose $(c, e, x) \approx (c', e', x')$. By definition, we can proceed from (c, e, x) to (c', e', x') by a sequence of steps

$$(c, e, x) = (c_0, e_0, x_0) \leftrightarrow (c_1, e_1, x_1) \leftrightarrow \cdots \leftrightarrow (c_k, e_k, x_k) = (c', e', x'). \quad \text{(A.15)}$$

Consider such a sequence with k minimal. All points in the sequence are then distinct, and each step consists of points either lying on the same A-curve (an "A-step") or on the same B-curve (a "B-step"). Moreover, if two successive steps are both A-steps or both B-steps, then we can contract them into a single step, contradicting minimality of k. Thus, the sequence must alternate between A-steps and B-steps. We will show that if $k \geq 3$, and either $c = c'$ or $e = e'$, then it is possible to replace some three consecutive steps of the sequence by two steps, again contradicting minimality of k. This will imply

$k = 1$ or 2. Note that if $c = c'$ or $e = e'$, then we cannot have $k = 2$ (because this would require the sequence to consists of one A-step and one B-step, leading to $c \neq c'$ and $e \neq e'$); thus we have $k = 1$, so (c', e', x') lies on the same A-curve or B-curve as (c, e, x), as needed.

Thus, assume $k \geq 3$. Assume also that $e = e'$ (the case $c = c'$ is similar). Since not all steps are A-steps, there must be some points (c_i, e_i, x_i) on the sequence with $e_i \neq e$. We can assume that there exist some points with $e_i > e$ (the case where some points have $e_i < e$ is similar). Let $\widehat{e}$ be the maximum value of e_i. Since the sequence begins and ends at $e < \widehat{e}$, there are some three consecutive steps

$$(c_i, e_i, x_i) \leftrightarrow (c_{i+1}, e_{i+1}, x_{i+1}) \leftrightarrow (c_{i+2}, e_{i+2}, x_{i+2}) \leftrightarrow (c_{i+3}, e_{i+3}, x_{i+3}) \tag{A.16}$$

with

$$e_{i+1} = e_{i+2} = \widehat{e}; \qquad e_i, e_{i+3} < \widehat{e}.$$

And of course

$$c_i = c_{i+1}, \qquad c_{i+2} = c_{i+3}.$$

Assume that $e_i \leq e_{i+3}$ (otherwise, the logic below applies after reversing the whole path (A.15)). The middle step of (A.16) is an A-step, telling us that

$$A(c_{i+1}, \widehat{e}, x_{i+1}) = A(c_{i+2}, \widehat{e}, x_{i+2}). \tag{A.17}$$

For each $e'' \in [e_{i+3}, \widehat{e}]$, consider whether the equality

$$A(c_{i+1}, e'', \xi^B(e''|c_{i+1}, \widehat{e}, x_{i+1})) = A(c_{i+2}, e'', \xi^B(e''|c_{i+2}, \widehat{e}, x_{i+2})) \tag{A.18}$$

holds. Note indeed that the ξ^B values in (A.18) are well-defined for each $e'' \in [e_{i+3}, \widehat{e}]$, since the ξ^B on the left side is well-defined at $e_i < e_{i+3}$ and at $\widehat{e}$ (where it equals x_i and x_{i+1} respectively) and the ξ^B on the right side is well-defined at e_{i+3} and at $\widehat{e}$ (where it equals x_{i+3} and x_{i+2} respectively). Moreover, (A.18) holds at $e'' = \widehat{e}$ since there it reduces to (A.17). Let $\widetilde{e}$ be the infimum of all values $e'' \in [e_{i+3}, \widehat{e}]$ for which (A.18) is satisfied; by continuity, it holds at $e'' = \widetilde{e}$. If $\widetilde{e} = e_{i+3}$ then we can shorten our three steps (A.16) to the two steps

$$(c_i, e_i, x_i) = (c_{i+1}, e_i, \xi^B(e_i|c_{i+1}, \widehat{e}, x_{i+1})) \leftrightarrow (c_{i+1}, e_{i+3}, \xi^B(e_{i+3}|c_{i+1}, \widehat{e}, x_{i+1}))$$

$$\leftrightarrow (c_{i+2}, e_{i+3}, \xi^B(e_{i+3}|c_{i+2}, \widehat{e}, x_{i+2})) = (c_{i+3}, e_{i+3}, x_{i+3}),$$

which is what we claimed. Thus, we now assume $\widetilde{e} > e_{i+3}$, and we will derive a contradiction.

Write

$$\widetilde{x}_1 = \xi^B(\widetilde{e}|c_{i+1}, \widehat{e}, x_{i+1}), \qquad \widetilde{x}_2 = \xi^B(\widetilde{e}|c_{i+2}, \widehat{e}, x_{i+2}) = \xi^A(c_{i+2}|c_{i+1}, \widetilde{e}, \widetilde{x}_1).$$

By assumption, each point on the A-curve between the points

$$(c_{i+1}, \widetilde{e}, \widetilde{x}_1) \leftrightarrow (c_{i+2}, \widetilde{e}, \widetilde{x}_2)$$

is contained in a local foliation leaf. By compactness, there exists a finite set of such local foliation leaves that cover the curve. In particular, because the curve is connected, we can form a sequence of points

$$(c_{i+1}, \widetilde{e}, \widetilde{x}_1) = (c'_0, \widetilde{e}, x'_0), (c'_1, \widetilde{e}, x'_1), \dots, (c'_l, \widetilde{e}, x'_l) = (c_{i+2}, \widetilde{e}, \widetilde{x}_2)$$

on the A-curve such that, for each j, the points $(c'_j, \widetilde{e}, x'_j)$ and $(c'_{j+1}, \widetilde{e}, x'_{j+1})$ lie in the same local foliation leaf S_j. By choosing $\epsilon > 0$ small enough, $\widetilde{e} - \epsilon$ also lies in the domain of each S_j. Therefore, by taking

$$x''_j = \xi^B(\widetilde{e} - \epsilon|c'_j, \widetilde{e}, x'_j)$$

(which is well-defined for ϵ small), we have

$$(c'_j, \widetilde{e} - \epsilon, x''_j), (c'_{j+1}, \widetilde{e} - \epsilon, x''_{j+1}) \in S_j.$$

Because S_j is a local foliation leaf, this implies

$$A(c'_j, \widetilde{e} - \epsilon, x''_j) = A(c'_{j+1}, \widetilde{e} - \epsilon, x''_{j+1}).$$

Stringing together these equalities for each j, we get

$$A(c'_0, \widetilde{e} - \epsilon, x''_0) = A(c'_l, \widetilde{e} - \epsilon, x''_l),$$

which is to say that (A.18) holds at $\widetilde{e} - \epsilon$. This is inconsistent with the minimality of $\widetilde{e}$, so we have our needed contradiction. □

Proof of Theorem 4.4. The theorem is now immediate by combining Lemmas A.3 and A.4. ☐

For the proof of Theorem 4.5, it will help to first define a partial order $\trianglelefteq$ on $\overline{C}\times\overline{E}\times\mathbb{R}_+$ by

$$(c,e,x)\trianglelefteq(c',e',x') \qquad \text{if} \qquad c\le c',\ e\ge e',\ \text{and } x\le x'.$$

We also write $(c,e,x)\triangleleft(c',e',x')$ if all three inequalities hold strictly.

Lemma A.5. *If an AB-foliation exists, then the transitive closure of the relation $\leftrightarrow$ is itself an AB-foliation.*[7]

Proof. Denote the given AB-foliation by $\approx'$. Denote the transitive closure of $\leftrightarrow$ by $\approx$. Evidently, $\approx$ is a (weakly) finer equivalence relation than $\approx'$, and if two points (c,e,x), (c',e',x') are on the same A-curve or B-curve then immediately $(c,e,x)\approx(c',e',x')$. Conversely, if (say) $e=e'$ and $(c,e,x)\approx(c',e',x')$, then also $(c,e,x)\approx'(c',e',x')$, hence $A(c,e,x)=A(c',e',x')$ by assumption that $\approx'$ is an AB-foliation; similarly if $c=c'$. ☐

Lemma A.6. *Suppose an AB-foliation exists, and let $\approx$ denote the transitive closure of $\leftrightarrow$. Then, there are no two distinct points $(c,e,x)\approx(c',e',x')$ with $(c',e',x')\triangleleft(c,e,x)$.*

Proof. Fix (c,e,x). Let S denote its equivalence class under $\approx$, and also write

$$S_-=\{(c',e')\mid(c',e',x')\in S \text{ for some } x'\}.$$

We claim that if $(c',e')\in S_-$ with $c'\le c$ and $e'\ge e$, then either $(c',e)\in S_-$ or $(c,e')\in S_-$. This will prove the lemma, since in the first case $(c',e',x')\approx(c',e,x'')\approx(c,e,x)$ for some x'', and the monotonicity of A-curves and B-curves then implies $x'\ge x''\ge x$, so we cannot have $(c',e',x')\triangleleft(c,e,x)$; and the second case is similar.

To prove the claim, since $(c',e')\in S_-$ implies $(c,e,x)\approx(c',e',x')$ for some x', we can proceed by "induction" on the number of $\leftrightarrow$ steps linking (c,e,x) to (c',e',x'). Thus, suppose $(c'',e'',x'')\leftrightarrow(c',e',x')$, and if $c''\le c$ and $e''\ge e$ then $(c'',e)\in S_-$ or $(c,e'')\in S_-$. Suppose $c'\le c$ and $e'\ge e$. We wish to show $(c',e)\in S_-$ or $(c,e')\in S_-$. Assume $c''=c'$ (the case $e''=e'$ is similar). If $e''<e$, then connectedness of the domain of B-curves implies $(c'',e)\in S_-$. If on the other hand $e''\ge e$, then we know $(c'',e)\in S_-$ or $(c,e'')\in S_-$.

[7] This is not a trivial equivalence, as the given foliation may be strictly coarser than the transitive closure of $\leftrightarrow$. For example, we can start from the latter foliation, take two equivalence classes whose ranges of c-values and e-values are disjoint, and merge them into a single equivalence class, thus creating a coarser AB-foliation.

In the first case we are done. In the second case, if $e'' \geq e'$ then using $(c, e) \in S_-$ and $(c, e'') \in S_-$, and connectedness of the domain of B-curves, we get $(c, e') \in S_-$. So suppose $e'' < e'$. We will show that $(c, e') \in S_-$ in this case too. Let $\widehat{x}$ be the value with $(c, e'', \widehat{x}) \in S$. Using

$$(c, e, x) \leftrightarrow (c, e'', \widehat{x}) \leftrightarrow (c'', e'', x'') \leftrightarrow (c'', e', x') = (c', e', x')$$

and $e \leq e'' < e'$, $c \geq c''$, by the monotonicity properties of A-curves and B-curves we have $x \leq \widehat{x} \leq x'' \leq x'$. This monotonicity also implies $B(c, e', \widehat{x}) \leq B(c, e'', \widehat{x})$. On the other hand, we claim $B(c, e', x') \geq B(c, e'', \widehat{x})$. If this is not the case, then $B(c, e', x' + \epsilon) < B(c, e'', \widehat{x})$ for small enough $\epsilon > 0$; on the other hand, $B(c, e'', x' + \epsilon) > B(c, e'', \widehat{x})$ since $x' + \epsilon > \widehat{x}$, so by continuity, there is some intermediate $e''' \in (e'', e')$ with $B(c, e''', x' + \epsilon) = B(c, e'', \widehat{x})$. Now, by connected domain and monotonicity of B-curves, we know there exists $x^\circ \in [x'', x']$ such that $(c'', e''', x^\circ) \in S$. Then, $(c'', e''', x^\circ) \approx (c, e''', x' + \epsilon)$, so these two points lie on an A-curve by Lemma A.5, so by monotonicity of A-curves, $x^\circ \geq x' + \epsilon$. But this contradicts $x^\circ \leq x'$. This contradiction gives the claim. So indeed $B(c, e', x') \geq B(c, e'', \widehat{x}) \geq B(c, e', \widehat{x})$. Thus, by continuity, there is some point $x^\dagger \in [\widehat{x}, x']$ with $B(c, e', x^\dagger) = B(c, e'', \widehat{x}) = B(c, e, x)$, i.e. $(c, e') \in S_-$ as needed. □

For the next lemma, and the proof of Theorem 4.5, define a curve $\gamma : (0, 1) \to \overline{C} \times \overline{E} \times \mathbb{R}_{++}$ by

$$\gamma(t) = \left(\underline{c} + t(\overline{c} - \underline{c}), \overline{e} - t(\overline{e} - \underline{e}), \frac{t}{1 - t} \right).$$

Also denote the coordinates of $\gamma(t)$ by $\gamma^c(t), \gamma^e(t), \gamma^x(t)$.

Lemma A.7. *Suppose an AB-foliation exists, and let $\approx$ denote the transitive closure of $\leftrightarrow$. Then, if S is an equivalence class of $\approx$, there is a unique value of t such that $\gamma(t) \in S$.*

Proof. Let $S_- = \{(c, e) \mid (c, e, x) \in S \text{ for some } x\}$. Then S_- is an open subset of $\overline{C} \times \overline{E}$. For each $(c, e) \in S_-$, there is a unique x such that $(c, e, x) \in S$ (by Lemma A.5), and there is a continuous function $\xi : S_- \to \mathbb{R}_{++}$ such that $S = \{(c, e, \xi(c, e)) \mid (c, e) \in S_-\}$. Moreover, ξ is weakly decreasing in c and increasing in e. (ξ can be defined locally by composition of ξ^A and ξ^B.)

We first claim that the set

$$T = \{t \mid (\gamma^c(t), \gamma^e(t)) \in S_-\}$$

is nonempty. To see this, pick any $(c^*, e^*) \in S_-$. Put $t^c = (c^* - \underline{c})/(\overline{c} - \underline{c}) = (\gamma^c)^{-1}(c^*)$ and

$t^e = (\overline{e} - e^*)/(\overline{e} - \underline{e}) = (\gamma^e)^{-1}(e^*)$. If $t^c = t^e$ then $(c^*, e^*) \in T$ already, so suppose $t^c < t^e$ (the case $t^c > t^e$ is similar, with the roles of c and e swapped and inequalities flipped).

Let R denote the rectangle $[\gamma^c(t^c), \gamma^c(t^e)] \times [\gamma^e(t^e), \gamma^e(t^c)]$. For $(c, e) \in R$, define $\chi(c, e) = (\gamma^c)^{-1}(c) - (\gamma^e)^{-1}(e)$. Let $\widetilde{T}$ denote the image of $R \cap S_-$ under χ. This is a relatively open subset of the interval $\chi(R) = [t^c - t^e, t^e - t^c]$, and it contains $\chi(c^*, e^*) = t^c - t^e$. If $0 \in \widetilde{T}$ then there exists t with $(\gamma^c(t), \gamma^e(t)) \in S_-$, as claimed. So suppose not. Then the component of $\widetilde{T}$ containing $t^c - t^e$ is an interval of the form $[t^c - t^e, \widetilde{t})$ for some $\widetilde{t} < 0$.

This means that there exists a sequence of points $(c_1, e_1), (c_2, e_2), \ldots$ in $R \cap S_-$ with $\chi(c_k, e_k)$ approaching $\widetilde{t}$ from below. By compactness of R, we can assume (c_k, e_k) converges to some limit $(\widetilde{c}, \widetilde{e}) \in R$, with $\chi(\widetilde{c}, \widetilde{e}) = \widetilde{t}$.

Pick an arbitrary $\widetilde{x} \in \mathbb{R}_{++}$. The foliation leaf containing $(\widetilde{c}, \widetilde{e}, \widetilde{x})$ gives a function $\widetilde{\xi}$, defined on an open rectangle $(\widetilde{c} - \epsilon, \widetilde{c} + \epsilon) \times (\widetilde{e} - \epsilon, \widetilde{e} + \epsilon)$ for some ϵ, such that $(c, e, \widetilde{\xi}(c, e)) \approx (\widetilde{c}, \widetilde{e}, \widetilde{x})$ for all (c, e) in that rectangle. In particular, for k sufficiently large, (c_k, e_k) lies in the domain of $\widetilde{\xi}$. Also, by again taking a subsequence, we can assume either (i) $\xi(c_k, e_k) \geq \widetilde{\xi}(c_k, e_k)$ for all k, or (ii) $\xi(c_k, e_k) < \widetilde{\xi}(c_k, e_k)$ for all k. (Because $(c_k, e_k) \in S_-$, we know $\xi(c_k, e_k)$ is well-defined.)

In case (i), for (c_k, e_k) sufficiently close to $(\widetilde{c}, \widetilde{e})$, we have $c_k < \widetilde{c} + \epsilon/2$ and $\chi(\widetilde{c} + \epsilon/2, e_k) > \widetilde{t}$. But then for every $c \in [c_k, \widetilde{c} + \epsilon/2]$, we have

$$A(c, e_k, \widetilde{\xi}(c, e_k)) = A(c_k, e_k, \widetilde{\xi}(c_k, e_k)) \leq A(c_k, e_k, \xi(c_k, e_k)) \leq A(c, e_k, \xi(c_k, e_k))$$

which implies that there is an intermediate value x' with $A(c, e_k, x') = A(c_k, e_k, \xi(c_k, e_k))$; that is, $(c, e_k) \in S_-$. If $\widetilde{c} + \epsilon/2 \leq \gamma^c(t^e)$ then every such point (c, e_k) is in $S_- \cap R$, so $\widetilde{T}$ contains the interval from $\chi(c_k, e_k) < \widetilde{t}$ to $\chi(\widetilde{c} + \epsilon/2, e_k) > \widetilde{t}$, contradicting the definition of $\widetilde{t}$. So $\widetilde{c} + \epsilon/2 > \gamma^c(t^e)$. But then still, every point (c, e_k) with $c_k \leq c \leq \gamma^c(t^e)$ is in $S_- \cap R$, so $\widetilde{T}$ contains the interval from $\chi(c_k, e_k)$ to $\chi(\gamma^c(t^e), e_k) \geq 0$ and thus contains 0, again contrary to assumption.

In case (ii), likewise, for every (c_k, e_k) sufficiently close to $(\widetilde{c}, \widetilde{e})$, we have $e_k < \widetilde{e} + \epsilon/2$ and $\chi(c_k, \widetilde{e} + \epsilon/2) > \widetilde{t}$. The same argument shows that if $\widetilde{e} + \epsilon/2 \leq \gamma^e(t^c)$ then every point (c_k, e) with $e_k \leq e \leq \widetilde{e} + \epsilon/2$ is in $S_- \cap R$, and thus $\widetilde{T}$ contains the interval from $\chi(c_k, e_k) < \widetilde{t}$ to $\chi(c_k, \widetilde{e} + \epsilon/2) > \widetilde{t}$, contradicting the definition of $\widetilde{t}$; and if $\widetilde{e} + \epsilon/2 > \gamma^e(t^c)$, then every point (c_k, e) with $e_k \leq e \leq \gamma^e(t^c)$ is in $S_- \cap R$, so $\widetilde{T}$ contains the interval from $\chi(c_k, e_k)$ to $\chi(c_k, \gamma^e(t^c)) \geq 0$ and so contains 0, again a contradiction.

Thus in each case we reach a contradiction, establishing the claim.

So T is a nonempty, open set. Let $(\underline{t}, \overline{t})$ be one of its connected components. Then $\xi(\gamma^c(t), \gamma^e(t))$ is continuous and weakly decreasing on $(\underline{t}, \overline{t})$. We claim that if $\overline{t} < 1$, then as $t \to \overline{t}$ from below, $\xi(\gamma^c(t), \gamma^e(t)) \to 0$. Suppose this is false; then $\xi(\gamma^c(t), \gamma^e(t))$ is bounded strictly above some positive number x°. By local foliation, there exists a continuous function ξ°, defined on a neighborhood of $(\gamma^c(\overline{t}), \gamma^e(\overline{t}))$, such that $(c, e, \xi^\circ(c, e)) \approx (\gamma^c(\overline{t}), \gamma^e(\overline{t}), x^\circ)$ throughout the neighborhood, and in particular $\xi^\circ(\gamma^c(\overline{t}), \gamma^e(\overline{t})) = x^\circ$. This implies that for t close to $\overline{t}$, $(\gamma^c(t), \gamma^e(t))$ is in this neighborhood, and $\xi(\gamma^c(t), \gamma^e(t)) > \xi^\circ(\gamma^c(t), \gamma^e(t))$. Then, by monotonicity of A-curves and the fact that A-curves do not cross, $(\gamma^c(\overline{t}), \gamma^e(\overline{t}))$ is in the domain of definition of ξ—that is, $(\gamma^c(\overline{t}), \gamma^e(\overline{t})) \in S_-$, contradicting $\overline{t} \notin T$. This proves the claim. A similar argument shows that, if $\underline{t} > 0$, then as $t \to \underline{t}$ from above, $\xi(\gamma^c(t), \gamma^e(t)) \to \infty$.

Thus, the function $\xi(\gamma^c(t), \gamma^e(t)) - \gamma^x(t)$ is defined and continuous on $(\underline{t}, \overline{t})$; for t close to $\overline{t}$ it becomes negative (either because $\overline{t} < 1$ and $\xi(\gamma^c(t), \gamma^e(t)) \to 0$, or because $\overline{t} = 1$ and $\gamma^x(t) \to \infty$); and for t close to $\underline{t}$ it becomes positive (for analogous reasons). Therefore, there is some intermediate t such that $\xi(\gamma^c(t), \gamma^e(t)) = \gamma^x(t)$. That is, $\gamma(t) \in S$, as required.

Finally, uniqueness of t is immediate from Lemma A.6.

□

Proof of Theorem 4.5. Suppose an AB-foliation exists; by Lemma A.5, we may assume it is the transitive closure of $\leftrightarrow$, and we denote it by $\approx$.

We will show that there exists a function U_{min}, defined on $\overline{C} \times \overline{E} \times \mathbb{R}_+$, such that

- U_{min} is constant on A-curves and B-curves;
- U_{min} is increasing in x; and
- for $x > 0$, $U_{min}(c, e, x)$ is continuously differentiable, with $\partial U_{min}/\partial x > 0$; and if Assumption 4.2 is satisfied, these properties extend to $x = 0$.

For $x > 0$, we can define $U_{min}(c, e, x)$ to be the unique value of t such that $(c, e, x) \approx \gamma(t)$, as given by Lemma A.7. We then complete the definition by defining $U_{min}(c, e, 0) = \inf_{x>0} U_{min}(c, e, x)$. We will show that this function meets the required conditions.

Fix any $(c^*, e^*, x^*) \in \overline{C} \times \overline{E} \times \mathbb{R}_{++}$; we will show the conditions are met on a neighborhood of (c^*, e^*, x^*). Put $(c^\dagger, e^\dagger, x^\dagger) = \gamma(U_{min}(c^*, e^*, x^*))$; thus there exists a sequence of distinct points

$$(c^*, e^*, x^*) = (c_0, e_0, x_0) \leftrightarrow (c_1, e_1, x_1) \leftrightarrow \cdots \leftrightarrow (c_k, e_k, x_k) = (c^\dagger, e^\dagger, x^\dagger).$$

For (c,e,x) in a neighborhood of (c^*,e^*,x^*), define functions $y_{-1},y_0,\dots,y_k$ of (c,e,x) by

$$\begin{aligned} y_{-1}(c,e,x) &= \xi^A(c^*|c,e,x); \\ y_0(c,e,x) &= \xi^B(e^*|c^*,e,y_{-1}(c,e,x)); \\ y_i(c,e,x) &= \begin{cases} \xi^A(c_i|c_{i-1},e_{i-1},y_{i-1}(c,e,x)) & \text{if } e_i = e_{i-1}, \\ \xi^B(e_i|c_{i-1},e_{i-1},y_{i-1}(c,e,x)) & \text{if } c_i = c_{i-1} \end{cases} \quad \text{for } i = 1,\dots,k. \end{aligned}$$

Thus we get

$$(c,e,x) \leftrightarrow (c_0,e,y_{-1}(c,e,x)) \leftrightarrow (c_0,e_0,y_0(c,e,x)) \leftrightarrow (c_1,e_1,y_1(c,e,x)) \leftrightarrow \cdots \leftrightarrow (c_k,e_k,y_k(c,e,x)).$$

Note that these functions are indeed defined, and continuously differentiable, for (c,e,x) in a neighborhood of (c^*,e^*,x^*). Moreover, using the implicit function theorem and the definition of ξ^A, one easily checks that, at any (c',c,e,x) where ξ^A is defined, $\partial\xi^A/\partial x > 0$; similarly, $\partial\xi^B/\partial x > 0$. Using these, it is an easy induction that $\partial y_i/\partial x > 0$.

Likewise, for (c,e,x) in a neighborhood of $(c^\dagger,e^\dagger,x^\dagger)$, we can define z_1,z_2 by

$$\begin{aligned} z_1(c,e,x) &= \xi^A(c^\dagger|c,e,x); \\ z_2(c,e,x) &= \xi^B(e^\dagger|c^\dagger,e,z_1(c,e,x)) \end{aligned}$$

so that

$$(c,e,x) \leftrightarrow (c^\dagger,e,z_1(c,e,x)) \leftrightarrow (c^\dagger,e^\dagger,z_2(c,e,x)).$$

A similar argument shows that these are continuously differentiable, with $\partial z_2/\partial x > 0$. Moreover, because $(c,e,x) \approx (c^\dagger,e^\dagger,z_2(c,e,x))$, the function z_2 is constant on A-curves and B-curves, hence $\partial z_2/\partial c \geq 0$ and $\partial z_2/\partial e \leq 0$. It follows that if we define $\widehat{z}(t) = z_2(\gamma(t))$, then $\widehat{z}$ is defined for t in a neighborhood of $U_{min}(c,e,x)$, and moreover is continuously differentiable there with $d\widehat{z}/dt > 0$.

This in turn means that $\widehat{z}$ is locally invertible, and its inverse has positive derivative. Finally, for (c,e,x) in a neighborhood of (c^*,e^*,x^*) we have

$$(c,e,x) \approx (c_k,e_k,y_k(c,e,x)) = (c^\dagger,e^\dagger,y_k(c,e,x)) \approx \gamma(\widehat{z}^{-1}(y_k(c,e,x)))$$

and therefore

$$U_{min}(c,e,x) = \widehat{z}^{-1}(y_k(c,e,x)).$$

Thus U_{min} is continuously differentiable, and because $d\widehat{z}/dt$ and $\partial y_k/\partial x$ are positive, so is $\partial U_{min}/\partial x$.

Also, as long as $x > 0$, U_{min} is constant on A-curves and B-curves, so $\partial U_{min}/\partial c \geq 0$ and $\partial U_{min}/\partial e \leq 0$.

This shows that U_{min} meets the required conditions for $x > 0$. Note also that since U is increasing in x for $x > 0$, we have $U(c, e, 0) = \lim_{x\downarrow 0} U(c, e, x)$. This, in turn, readily implies that U_{min} is constant on A-curves and B-curves and is increasing in x, even including points with $x = 0$.

Finally, if Assumption 4.2 is satisfied, we modify the construction as follows: let $\widetilde{A}$ be the function given in the assumption, with $A(c, e, x) = \widetilde{A}(c, e, x + \alpha)$; perform the preceding construction with $\widetilde{A}$ in place of A, obtaining a function $\widetilde{U}_{min}(c, e, x)$; then define $U_{min}(c, e, x) = \widetilde{U}_{min}(c, e, x + \alpha)$. It is immediate that the desired properties hold for U_{min} for all $x \geq 0$.

Now we can check that the welfarist criterion $\mathcal{W}_{min}$ defined by $U = U_{min}$ and $W(u_1, \ldots, u_n) = \min\{u_1, \ldots, u_n\}$ represents equal opportunity for one-dimensional populations, and that under Assumption 4.2, this criterion is regular. The latter property is immediate from what we have already established.

Consider any population $P = C \times E$, with $E = \{e\}$, and any total quantity $\overline{x}$. From Lemma 4.3, there is a unique allocation X^* maximizing $\mathcal{V}^{EOp}$, characterized by a value A^*: every agent (c, e) with $A(c, e, 0) < A^*$ is given $X^*(c, e)$ so that $A(c, e, X^*(c, e)) = A^*$, and every agent (c, e) with $A(c, e, 0) \geq A^*$ is given 0. Moreover, because A is increasing in c, the latter set of agents (if nonempty) is characterized by $c \geq c^*$ for some cutoff c^*. Then, because U_{min} is constant on A-curves and weakly increasing in c, $U_{min}(c, e, X^*(c, e))$ is constant (at some value U^*_{min}) for the former set of agents, and takes values $\geq U^*_{min}$ for the latter set. Therefore $\mathcal{W}_{min}(X^*) = U^*_{min}$. Any other allocation X' would have to give some agent in the former set less than X^* does, and therefore $U_{min}(c, e, X'(c, e)) < U^*_{min}$, so $\mathcal{W}_{min}(X') < U^*_{min}$. Thus, X^* is the unique maximizer of $\mathcal{W}_{min}$, as required. A similar argument applies for populations with $|C| = 1$ instead of $|E| = 1$.

Finally, to construct the representation using the sum aggregator, let $Z : [0, 1] \to \mathbb{R}_{++}$ be any continuously differentiable function with $Z' < 0$. (For example, $Z(t) = 2 - t$ works.) Define

$$U_{sum}(c, e, x) = \int_0^x Z(U_{min}(c, e, \widetilde{x}))\, d\widetilde{x}.$$

We will check that the welfarist criterion $\mathcal{W}_{sum}$, defined by $U = U_{sum}$ and $W(u_1, \ldots, u_n) = u_1 + \cdots + u_n$, again represents equal opportunity for one-dimensional populations, and is regular under Assumption 4.2.

It is clear that this is indeed a welfarist criterion, i.e. U_{sum} is strictly increasing in x—indeed, differentiable in x, with $\partial U_{sum}/\partial x = Z(U_{min}(c, e, x)) > 0$. Under Assumption 4.2, the stronger continuous differentiability requirement for regularity follows as a direct consequence of the continuous differentiability of U_{min} and Leibniz's rule for differentiation under the integral.

To see that $\mathcal{W}_{sum}$ represents equal opportunity for one-dimensional populations, consider a distribution problem $(P, \overline{x})$ with P one-dimensional. U_{sum} is strictly concave in x, so $\mathcal{W}_{sum}$ is strictly concave, and therefore its maximizer X^* in the given distribution problem (which exists by continuity) is unique. Now logic identical to that of Lemma 4.3 (part 2) implies that there is some U^*_{min} such that every agent with $U_{min}(c, e) \leq U^*_{min}$ is given a quantity so that $U_{min}(c, e, X^*(c, e)) = U^*_{min}$, and every agent with $U_{min}(c, e) > U^*_{min}$ is given quantity 0. But this is exactly the condition for X^* to maximize $\mathcal{W}_{min}$; thus $\mathcal{W}_{sum}$ selects the same allocations as $\mathcal{W}_{min}$ does. □

A.4 Proofs from Section 5

Proof of Proposition 5.1. Let $\mathcal{W}_{sum}$ and U_{sum} be as in the proof of Theorem 4.5. It suffices to show that the canonical allocation is optimal for any regular $\mathcal{W}$ with a weakly separable aggregator; its uniqueness then follows from the fact that $\mathcal{W}_{sum}$ is such an aggregator and is strictly concave over $\Delta_{\overline{x}}(P)$, so its maximizer is unique.

For any two allocations $X, X' \in \Delta_{\overline{x}}(P)$, say that X' is a *feasible reallocation from X* if either

- there exists $c \in C$ such that the restriction of X' to $\{c\} \times E$ equals the equal-opportunity choice for problem $(\{c\} \times E, \sum_e X(c, e))$, and X' coincides with X on $(C \setminus \{c\}) \times E$; or

- there exists $e \in E$ such that the restriction of X' to $C \times \{e\}$ equals the equal-opportunity choice for problem $(C \times \{e\}, \sum_c X(c, e))$, and X' coincides with X on $C \times (E \setminus \{e\})$.

(Thus, the feasible reallocations from X are those reached by choosing an element of C or E, and reallocating canonically within the corresponding one-dimensional subpopulation.) Because the reallocation increases the value of $\mathcal{W}$ within the subpopulation, weak separability implies that $\mathcal{W}(X') \geq \mathcal{W}(X)$.

Now start from any allocation $X \in \Delta_{\overline{x}}(P)$. Define a sequence of allocations $X = X^1, X^2, \ldots$ as follows: given X^k, we consider, among all feasible reallocations from X^k, one that maximizes the value of $\mathcal{W}_{sum}$, and take X^{k+1} to be this reallocation. Evidently, both $\mathcal{W}(X^k)$ and $\mathcal{W}_{sum}(X^k)$ are weakly increasing in k. By compactness, there exists a subsequence of (X^k) that converges to a limit X^∞, and then by continuity $\mathcal{W}_{sum}(X^k) \to \mathcal{W}_{sum}(X^\infty)$ and $\mathcal{W}(X^k) \to \mathcal{W}(X^\infty)$ also. In particular, $\mathcal{W}(X) \leq \mathcal{W}(X^\infty)$.

We claim that X^∞ must equal the canonical allocation X^*, which will suffice to show that X^* indeed maximizes $\mathcal{W}$. Note that within each one-dimensional subpopulation $\{c\} \times E$ or $C \times \{e\}$, X^∞ must coincide with the equal-opportunity choice (for the total quantity allocated to that subpopulation by X^∞): otherwise, we could reallocate within the subpopulation so as to strictly increase the value of $\mathcal{W}_{sum}$, say to $\mathcal{W}_{sum}(X^\infty) + \epsilon$; but then, for large enough k, the corresponding reallocation from X^k would increase the value of $\mathcal{W}_{sum}$ to at least $\mathcal{W}_{sum}(X^\infty) + \epsilon/2 > \mathcal{W}_{sum}(X^{k+1})$, contradicting the fact that X^{k+1} was chosen among feasible reallocations from X^k to maximize the value of $\mathcal{W}_{sum}$.

Now, both X^∞ and X^* have the property that, within each one-dimensional subpopulation, they coincide with the equal-opportunity choice. We will show that this property forces X^∞ and X^* to be identical. Let $c_* = \min(C)$ and $e^* = \max(E)$, and assume that $X^\infty(c_*, e^*) \geq X^*(c_*, e^*)$ (the argument when $X^\infty(c_*, e^*) \leq X^*(c_*, e^*)$ is similar). If $X^\infty(c_*, e^*) = 0$, then $X^*(c_*, e^*) = 0$, and (by the monotonicity of B-curves) $X^\infty(c_*, e) = X^*(c_*, e) = 0$ for all e. Otherwise, $(c_*, e^*, X^\infty(c_*, e^*))$ lies on a (weakly) higher B-curve than $(c_*, e^*, X^*(c_*, e^*))$, hence $X^\infty(c_*, e) \geq X^*(c_*, e)$ for all e. Now, for each e, repeating the foregoing argument in the c-dimension gives $X^\infty(c, e) \geq X^*(c, e)$ for each c. Since both allocations have the same total quantity $\overline{x}$, they must coincide. □

Proof of Lemma 5.2. First consider $\mathcal{V}^{\Sigma m}$. Suppose X maximizes $\mathcal{V}^{\Sigma m}$ over $\Delta_P(\overline{x})$. For each c and e, if $A(c, e, X(c, e)) \neq \min_{c'} A(c', e, X(c', e))$ then $X(c, e) = 0$: otherwise, we could reduce $X(c, e)$ by some $\epsilon > 0$, increase $X(c', e)$ by $\frac{\epsilon}{|C|-1}$ for each $c' \neq c$, and thereby improve the objective.

Also, for each e, the quantity $m_e(X) = \min_c A(c, e, X(c, e))$ is a weakly concave functional of the allocation X. Now, suppose X, X' are two allocations that both maximize $\mathcal{V}^{\Sigma m}$, and put $X'' = (X + X')/2$. We thus have

$$\mathcal{V}^{\Sigma m}(X'') = \sum_e m_e(X'') \geq \sum_e \frac{1}{2}\left(m_e(X) + m_e(X')\right) = \frac{1}{2}\left(\mathcal{V}^{\Sigma m}(X) + \mathcal{V}^{\Sigma m}(X')\right), \tag{A.19}$$

and by weak concavity we must have equality throughout, and X'' also maximises $\mathcal{V}^{\Sigma m}$.

Hence, for each e, and any c such that $A(c,e,X''(c,e)) = m_e(X'')$, we must have $X(c,e) = X'(c,e) = X''(c,e)$: otherwise, we would have $A(c,e,X''(c,e)) > \frac{1}{2}\left(A(c,e,X(c,e)) + A(c,e,X'(c,e))\right)$ strictly, leading to strict inequality in the m_e-comparisons in (A.19), a contradiction. Then $m_e(X), m_e(X') \leq m_e(X'')$, and we must have equality. Moreover, for any c such that $A(c,e,X''(c,e)) > m_e(X'')$, we must have $X''(c,e) = 0$ by the previous paragraph. Then immediately $X(c,e) = X'(c,e) = 0$. Thus, X and X' agree everywhere, so the maximizer of $\mathcal{V}^{\Sigma m}$ is unique.

Now consider $\mathcal{V}^{m\Sigma}$. For each c, the value $s_c(X) = \sum_e A(c,e,X(c,e))$ is a strictly concave functional of the partial allocation $(X(c,\cdot))$. Suppose that X and X' both maximize $\mathcal{V}^{m\Sigma}$, and put $X'' = (X + X')/2$. Note that $\mathcal{V}^{m\Sigma}$ is a weakly concave function of the allocation, so X'' must also maximize $\mathcal{V}^{m\Sigma}$. For any c such that $s_c(X'') = \min_{c'} s_{c'}(X'')$, if $X(c,e) \neq X'(c,e)$ for some e, then $s_c(X'') > \frac{1}{2}(s_c(X) + s_c(X'))$ by strict concavity, implying $\mathcal{V}^{m\Sigma}(X'') > \frac{1}{2}\left(\mathcal{V}^{m\Sigma}(X) + \mathcal{V}^{m\Sigma}(X')\right)$, contradicting that X and X' are maximizers. Thus, for each such c, we have $X(c,e) = X'(c,e)$ for all e. Moreover, for any c such that $s_c(X'') > \min_{c'} s_{c'}(X'')$, we must have $X''(c,e) = 0$ for every e, since otherwise we could modify X'' by giving ϵ less of the good to agent (c,e) and giving $\frac{\epsilon}{|C|-1}$ to agent (c',e) for each $c' \neq c$, thereby increasing the objective and contradicting optimality of X''. This again implies $X(c,e) = X'(c,e) = 0$ for each such c and all e. So X and X' coincide. □

Proof of Proposition 5.3. For brevity, denote the three allocations by $X^*, X^{\Sigma m}$, $X^{m\Sigma}$ (suppressing the $(P, \overline{x})$). Suppose first that $X^*(c,e) > 0$ for all $(c,e) \in P = C \times E$. Then, for every e, the advantage level $A(c,e,X^*(c,e))$ is constant across $c \in C$; hence, the sum $\sum_e A(c,e,X^*(c,e))$ is also constant across c, and we denote its common value by A^{**}. Suppose that $X^{m\Sigma}$ is different from X^*. Then the optimal min-of-sums value is some $\widetilde{A} > A^{**}$. However, for each fixed c, if we consider the distribution problem $(\{c\} \times E, \sum_e X^*(c,e))$, the allocation that maximizes total advantage is exactly given by the restriction of X^* to this one-dimensional population (because the corresponding points lie on the same B-curve, by construction). Since $X^{m\Sigma}$ achieves a higher total advantage than X^* on this subpopulation, we must then have $\sum_e X^{m\Sigma}(c,e) > \sum_e X^*(c,e)$. Summing over each $c \in C$, we then have

$$\overline{x} = \sum_{c,e} X^{m\Sigma}(c,e) > \sum_{c,e} X^*(c,e) = \overline{x},$$

a contradiction. Thus $X^{m\Sigma} = X^*$.

Suppose that $X^{\Sigma m}$ is different from X^*. For each e, the difference $X^{\Sigma m}(c,e)-X^*(c,e)$ must be positive for all c, negative for all c, or zero for all c. To see this, suppose that $X^{\Sigma m}(c,e) \le X^*(c,e)$ but $X^*(c',e) \le X^{\Sigma m}(c',e)$, with at least one inequality strict. Then

$$A(c,e,X^{\Sigma m}(c,e)) \le A(c,e,X^*(c,e)) = A(c',e,X^*(c',e)) \le A(c',e,X^{\Sigma m}(c',e)),$$

with one inequality strict, and $X^{\Sigma m}(c',e) > 0$. Then, as in the proof of Lemma 5.2, we could start from allocation $X^{\Sigma m}$, take some of the good from agent (c',e) and redistribute so as to improve the sum-of-mins objective, a contradiction. Now, since $X^{\Sigma m} \neq X^*$ but both allocations use the same total quantity $\overline{x}$, this implies there must be some e such that the difference $X^{\Sigma m}(c,e) - X^*(c,e)$ is negative for all c, and some e such that it is positive for all c; call these e_- and e_+. Then for each c, we have

$$\frac{\partial A}{\partial x}(c,e_-,X^{\Sigma m}(c,e_-)) > \frac{\partial A}{\partial x}(c,e_-,X^*(c,e_-)) = \frac{\partial A}{\partial x}(c,e_+,X^*(c,e_+)) > \frac{\partial A}{\partial x}(c,e_+,X^{\Sigma m}(c,e_+)).$$

Consequently, for small $\epsilon > 0$, if we start from $X^{\Sigma m}$, and reallocate the good so as to increase the advantage enjoyed by agent (c,e_-) by amount ϵ and decrease the advantage of agent (c,e_+) by ϵ (for each $c \in c$), the total quantity of the good used decreases, while the value of the sum-of-mins objective stays constant. We can then redistribute the remaining supply of the good so as to increase the sum-of-mins objective, contradicting optimality of $X^{\Sigma m}$. Thus $X^{\Sigma m} = X^*$ also.

Next, start from the assumption that $X^{m\Sigma}(c,e) > 0$ for all (c,e). Then, again, the sum $\sum_e A(c,e,X^{m\Sigma}(c,e))$ must be constant across $c \in C$ (otherwise we could redistribute so as to improve the objective). Moreover, for any given c, the points $(c,e,X^{m\Sigma}(c,e))$, as e varies, must all lie on the same B-curve (otherwise we could again improve the objective) and therefore all lie on the same leaf of the AB-foliation. Consequently, for each c, the difference $X^*(c,e) - X^{m\Sigma}(c,e)$ must be negative for all e, positive for all e, or zero for all e: if (for example) $X^*(c,e) > X^{m\Sigma}(c,e)$, then both X^* and $X^{m\Sigma}$ place any other agent (c,e') on the same foliation leaf as (c,e), and so $X^*(c,e') > X^{m\Sigma}(c,e')$; similarly for the $X^*(c,e) = X^{m\Sigma}(c,e)$ case. So, if $X^* \neq X^{m\Sigma}$, then there must exist some value c such that $X^*(c,e) - X^{m\Sigma}(c,e)$ is negative for all e, and another such that it is positive for all e; call these values c_- and c_+. Then we have

$$A(c_-,e,X^{m\Sigma}(c_-,e)) > A(c_-,e,X^*(c_-,e)) \ge A(c_+,e,X^*(c_+,e)) > A(c_+,e,X^{m\Sigma}(c_+,e))$$

for each e. (The middle inequality reflects the fact that X^* puts (c_-, e) and (c_+, e) on the same foliation sheet, except for the possible corner case $X^*(c_-, e) = 0$; note that the opposite corner case $X^*(c_+, e) = 0$ cannot occur due to $X^*(c_+, e) > X^{m\Sigma}(c_+, e)$.) Consequently, we cannot have $\sum_e A(c_-, e, X^{m\Sigma}(c_-, e)) = \sum_e A(c_+, e, X^{m\Sigma}(c_+, e))$ as claimed. So, $X^* = X^{m\Sigma}$, and now we are in the previous case where X^* gives positive quantities to all agents.

Finally, start from the assumption that $X^{\Sigma m}(c, e) > 0$ for all (c, e). For each e, $A(c, e, X^{\Sigma m}(c, e))$ must be constant across all c (otherwise we could redistribute and improve the objective); thus the points $(c, e, X^{\Sigma m})$ lie on the same A-curve. Therefore, for each e, the difference $X^*(c, e) - X^{\Sigma m}(c, e)$ is either positive for all c, zero for all c, or negative for all c; the argument is the same as in the previous paragraph. If $X^* \neq X^{\Sigma m}$, then there is some e_- such that $X^*(c, e_-) < X^{\Sigma m}(c, e_-)$ for all c, and another e_+ such that $X^*(c, e_+) > X^{\Sigma m}(c, e_+)$ for all c. Then, for each c, we have

$$\frac{\partial A}{\partial x}(c, e_+, X^{\Sigma m}(c, e_+)) > \frac{\partial A}{\partial x}(c, e_+, X^*(c, e_+)) \geq \frac{\partial A}{\partial x}(c, e_-, X^*(c, e_-)) > \frac{\partial A}{\partial x}(c, e_-, X^{\Sigma m}(c, e_-))$$

where the middle inequality comes from X^* putting (c, e_+) and (c, e_-) on the same B-curve, except in the corner case $X^*(c, e_-) = 0$. Therefore, for small $\epsilon > 0$, we can start from $X^{\Sigma m}$ and reallocate the good so as to increase the advantage of agent (c, e_+) by ϵ and decrease the advantage of (c, e_-) by ϵ for each c, reducing the total quantity allocated. Then we can further redistribute the excess supply to all agents. This increases the sum-of-mins objective, contradicting optimality of $X^{\Sigma m}$. So $X^* = X^{\Sigma m}$, and we are again now in the case where $X^*(c, e) > 0$ for all (c, e). □

Proof of Lemma 5.4. First, note that U is weakly increasing in c and weakly decreasing in e. (This was shown in the proof of Theorem 4.5.)

Now suppose the flat-curve condition is satisfied, and fix u^*. We show that there exist $\widehat{c}, \widehat{e}$ satisfying the requirements of the product-domain condition. Assume there are some pairs $(c, e) \in \overline{C} \times \overline{E}$ such that $U(c, e, 0) < u^*$ (otherwise we can take $\widehat{c} = \underline{c}$, $\widehat{e} = \overline{e}$).

Let $\widehat{c}$ be the supremum of c over all such pairs (c, e), and let $\widehat{e}$ be the infimum of e over all such pairs. So it is immediate that, whenever $c > \widehat{c}$ or $e < \widehat{e}$, we have $U(c, e, 0) \geq u^*$, and this inequality extends to pairs with $c = \widehat{c}$ or $e = \widehat{e}$ by continuity (note that if $\widehat{c} = \overline{c}$ or $\widehat{e} = \underline{e}$ this statement is vacuous). Conversely, consider (c, e) with $U(c, e, 0) \geq u^*$; we need to show that $c \geq \widehat{c}$ or $e \leq \widehat{e}$. By the flat-curve condition (and the fact that U is constant on A-curves and B-curves), we have $U(c', e, 0) = U(c, e, 0)$ for all $c' < c$ or $U(c, e', 0) = U(c, e, 0)$ for all $e' > e$; assume the first case. Then, by monotonicity in c,

$U(c', e, 0) \geq u^*$ for all $c' > c$ as well. Then by monotonicity in e, $U(c', e', 0) \geq u^*$ for all c' and all $e' \leq e$. But then the definition of $\widehat{e}$ implies that $\widehat{e} \geq e$ as needed. Similarly, the second case ($U(c, e', 0) = U(c, e, 0)$ for all $e' > e$) implies that $\widehat{c} \leq c$.

Conversely, suppose the product-domain condition is satisfied. For any (c, e), consider $u^* = U(c, e, 0)$. Then, by the product-domain condition, either $c \geq \widehat{c}(u^*)$ or $e \leq \widehat{e}(u^*)$. The first case implies $U(c, e', 0) \geq u^* = U(c, e, 0)$ for all $e' > e$ (by definition of $\widehat{c}(u^*)$), so we have equality by monotonicity of U. Thus, for all $e' > e$, $(c, e', 0)$ and $(c, e, 0)$ lie on the same B-curve. Likewise, the second case implies that for all $c' < c$, $(c', e, 0)$ and $(c, e, 0)$ lie on the same A-curve. □

Proof of Proposition 5.5. First, suppose the flat-curve condition is satisfied. Let U index the foliation leaves. Fix a distribution problem $(P, \overline{x})$ with $P = C \times E$, and write $X^* = X^*(P, \overline{x})$. From Lemma 5.4 and the definition of X^*, there are some thresholds $\widehat{c}, \widehat{e}$ and some value U^* such that for each agent (c, e) with $c < \widehat{c}$ and $e > \widehat{e}$, we have $X^*(c, e) > 0$ and $U(c, e, X^*(c, e)) = U^*$; and for each other agent, $X^*(c, e) = 0$ and $U(c, e, 0) \geq U^*$. We can assume some agent $(c, e) \in P$ has $c < \widehat{c}$ and $e > \widehat{e}$, otherwise we are in the trivial case $\overline{x} = 0$.

We check that for each $e \in E$, the value of $A(c, e, X^*(c, e))$ is constant across all $c \in C$ with $c < \widehat{c}$. For $e > \widehat{e}$ this is immediate from the fact that X^* places all such agents (c, e) on the same A-curve. For $e \leq \widehat{e}$, the claim is that $A(c, e, 0)$ is constant across all $c < \widehat{c}$, $c \in C$. If not, there exist $c' < c''$ in C, both less than $\widehat{c}$, where $A(c', e, 0) \neq A(c'', e, 0)$; then, by the flat-curve condition, $B(c'', e, 0) = B(c'', e', 0)$ for all $e' > e$. But this contradicts that $U(c'', e, 0) \geq U^*$ while $U(c'', e', 0) < U^*$ for $e' = \max(E) > \widehat{e}$.

It follows that the value of $\sum_e A(c, e, X^*(c, e))$ is constant across all $c \in C$ with $c < \widehat{c}$; call this value A^{**}. And since the canonical allocation puts agents with higher c-values at a weakly higher advantage level, we also have $\sum_e A(c, e, X^*(c, e)) \geq A^{**}$ for $c \geq \widehat{c}$, so that A^{**} is the min-of-sums value for X^*. We claim that A^{**} is in fact the optimal min-of-sums value, so that the min-of-sums allocation $X^{m\Sigma}$ coincides with X^*. Suppose instead that $X^{m\Sigma}$ attains a min-of-sums value $\widetilde{A} > A^{**}$. Then, for each $c < \widehat{c}$, we must have $\sum_e X^{m\Sigma}(c, e) > \sum_e X^*(c, e)$, as in the proof of Proposition 5.3. And for $c \geq \widehat{c}$ we have $\sum_e X^{m\Sigma}(c, e) \geq \sum_e X^*(c, e)$ since the right side is zero. So $\overline{x} = \sum_{c,e} X^{m\Sigma}(c, e) > \sum_{c,e} X^*(c, e) = \overline{x}$, a contradiction.

Next, we must show that the sum-of-mins allocation $X^{\Sigma m}$ also coincides with X^*. Suppose not. Then there exist some agent (c_-, e_-) who receives strictly less under $X^{\Sigma m}$ than X^*, and some agent (c_+, e_+) who receives more under $X^{\Sigma m}$ than X^*. In $X^{\Sigma m}$, any two agents with the same e-value are put on the same A-curve (aside from corners with $x = 0$), as in the proof of Lemma 5.2—as is the case for X^*. Therefore $X^{\Sigma m}(c, e_-) \leq X^*(c, e_-)$ for all c, and $X^{\Sigma m}(c, e_+) \geq X^*(c, e_+)$ for all c. Furthermore, $X^{\Sigma m}(c_-, e_-) < X^*(c_-, e_-)$ implies $e_- > \widehat{e}$.

We claim that $X^{\Sigma m}(c, e_+) > X^*(c, e_+)$ for all $c < \widehat{c}$. Indeed, if $e_+ > \widehat{e}$ then this claim follows from $X^{\Sigma m}(c, e_+) \geq X^*(c, e_+)$, and the strict inequality comes from the fact that both $X^{\Sigma m}$ and X^* place all agents (c, e_+) on the same A-curve (and agents with $c < \widehat{c}$ are not at corners) and the fact that the strict inequality holds for $c = c_+$. Otherwise, the claim follows from the fact that all for all $c < \widehat{c}$, $(c, \widehat{e}, 0)$ lie on the same A-curve (as shown above), and $X^{\Sigma m}$ also places all such $(c, \widehat{e})$ on the same A-curve, and the latter curve lies strictly above zero (since $X^{\Sigma m}(c_+, e_+) > 0$ implies $X^{\Sigma m}(\min(C), e_+) > 0$ by monotonicity of A-curves).

For each $c < \widehat{c}$, we have

$$\frac{\partial A}{\partial x}(c, e_-, X^{\Sigma m}(c, e_-)) \geq \frac{\partial A}{\partial x}(c, e_-, X^*(c, e_-)) \geq \frac{\partial A}{\partial x}(c, e_+, X^*(c, e_+)) > \frac{\partial A}{\partial x}(c, e_+, X^{\Sigma m}(c, e_+))$$

where the strictness of the last inequality holds by the preceding paragraph. Therefore, for small $\epsilon > 0$, we can modify $X^{\Sigma m}$ by increasing the advantage of agent (c, e_-) by ϵ and decreasing the advantage of agent (c, e_+) by ϵ for each $c < \widehat{c}$ (note that the latter agents all receive positive quantities in $X^{\Sigma m}$ by the previous paragraph). This change strictly reduces the total amount of the good used. The minimum advantage among agents with $e = e_+$ is reduced by no more than ϵ, while the minimum advantage among agents with $e = e_-$ is increased by exactly ϵ as long as we check that, at $X^{\Sigma m}$, this minimum was not attained by any agent with $c \geq \widehat{c}$ (and as long as ϵ is small enough). However, we know that any such agent (c, e_-) is placed at the $x = 0$ corner by both $X^{\Sigma m}$ and X^*, and also X^* assigns a positive (non-corner) quantity to agent (c_-, e_-); hence,

$$A(c, e_-, X^{\Sigma m}(c, e_-)) = A(c, e_-, X^*(c, e_-)) \geq A(c_-, e_-, X^*(c_-, e_-)) > A(c_-, e_-, X^{\Sigma m}(c_-, e_-)),$$

showing the needed non-minimality for (c, e_-).

Thus the change from $X^{\Sigma m}$ weakly increases the value of the sum-of-mins objective, while using strictly less of the good. The excess of the good can then be redistributed to strictly increase the objective, contradicting optimality of $X^{\Sigma m}$. This contradiction shows that indeed $X^{\Sigma m} = X^*$.

Now we prove the second part of the proposition. Suppose the flat-curve condition is not satisfied, at some $(c,e) \in \overline{C} \times \overline{E}$. Then there exists $c' < c$ and $e' > e$ such that $A(c',e,0) < A(c,e,0)$ and $B(c,e',0) < B(c,e,0)$. So, by choosing U^* slightly less than $U(c,e,0)$, we can find an allocation X^* on the population $P = \{c,c'\} \times \{e,e'\}$ that gives positive quantities to agents (c',e), (c,e'), and (c',e'), with $U(c',e,X^*(c',e)) = U(c,e',X^*(c,e')) = U(c',e',X^*(c',e')) = U^*$, while $X^*(c,e) = 0$. Thus, X^* is a canonical allocation in a problem $(P,\overline{x})$.

We show that X^* does not maximize either the min-of-sums or sum-of-mins objective for this problem, so that the corresponding optimal allocations must be different from X^*. For the min-of-sums, note that $A(c,e,X^*(c,e)) > A(c',e,X^*(c',e))$ while $A(c,e',X^*(c,e')) = A(c',e',X^*(c',e'))$, thus the sum of A-values across $\{e,e'\}$ is lower at c' than at c, so we can strictly improve it by taking some of the good from agent (c,e') and giving it to (c',e'). For the sum-of-mins, note that $A(c',e,X^*(c',e)) < A(c,e,X^*(c,e))$ and $\frac{\partial A}{\partial x}(c',e,X^*(c',e)) = \frac{\partial A}{\partial x}(c',e',X^*(c',e'))$, so if we consider perturbing X^* by increasing the amount given to agent (c',e) by $\epsilon > 0$ and decreasing the amounts of agents (c',e') and (c,e') by $\lambda\epsilon$ and $(1-\lambda)\epsilon$ (where λ is a fixed value slightly less than 1), the sum-of-mins objective then equals the sum of the advantages of agents (c',e) and (c',e'), whose derivative with respect to ϵ at $\epsilon = 0$ is $\frac{\partial A}{\partial x}(c',e,X^*(c',e)) - \lambda\frac{\partial A}{\partial x}(c',e',X^*(c',e')) = (1-\lambda)\frac{\partial A}{\partial x}(c',e,X^*(c',e)) > 0$. So, some such perturbation improves the objective.

Finally, we need to show that the min-of-sums and sum-of-mins allocations for this problem also do not coincide. Suppose that some allocation X' does optimize both these objectives simultaneously. To be sum-of-mins optimal, X' must put (c',e) and (c,e) on the same A-curve modulo corners, and likewise for (c',e') and (c,e'). Then $A(c',e,X'(c,e)) + A(c',e',X'(c',e')) \leq A(c,e,X'(c,e)) + A(c,e',X'(c,e'))$, hence (c',e) and (c',e') must be on the same B-curve modulo corners, otherwise we could improve the min-of-sums. Consequently, if we set $U^{**} = U(c',e',X'(c',e'))$, then X' either puts agent (c,e') on the same foliation leaf U^{**}, or gives this agent zero and is then already above that foliation leaf. Likewise, the same is true for agent (c',e), and then for (c,e). In conclusion, X' is the canonical allocation for population P and some quantity $\overline{x}'$. Since we must have $\overline{x}' = \overline{x}$, X' coincides with X^*. But we already showed this cannot happen. □

Proof of Proposition 5.6. It suffices to show that every point $(c^*,e^*,x^*) \in \overline{C} \times \overline{E} \times \mathbb{R}_{++}$ is contained in some local foliation leaf, since the result then follows from Lemma A.4.

Fix (c^*,e^*,x^*), and define the functions x_1,x_2,x_3,x_4 as in the proof of Lemma A.3. As in that proof, we can choose neighborhoods $\overline{C}',\overline{E}'$ of c^*,e^* respectively, so that each x_j is defined throughout $\overline{C}' \times \overline{E}'$. And, as in that proof, if $x_4(c,e) = x^*$ for all $(c,e) \in \overline{C}' \times \overline{E}'$, then x_2 locally respects AB-foliation, and we are done. So, suppose there exists some (c, e) such that $x_4(c, e) \neq x^*$, and seek a contradiction.

Consider the distribution problem with population $P = \{c^*, c\} \times \{e^*, e\}$ and total quantity

$$\overline{x} = x^* + x_1(c, e) + x_2(c, e) + x_3(c, e).$$

Consider the allocation

$$X^* = X^{\Sigma m}(P, \overline{x}) = X^{m\Sigma}(P, \overline{x})$$

(by assumption, these two optimal allocations coincide). Compare it to the allocation $\widehat{X}$ given by

$$\widehat{X}(c^*, e^*) = x^*; \qquad \widehat{X}(c, e^*) = x_1(c, e); \qquad \widehat{X}(c, e) = x_2(c, e); \qquad \widehat{X}(c^*, e) = x_3(c, e).$$

Note that if X^* allocates weakly more than $\widehat{X}$ to some agent (c', e'), then the other agent with the same e-value, (c'', e'), must be placed on the same A-curve as (c', e')—and therefore also receives weakly more under X^* than under $\widehat{X}$ (since $\widehat{X}$ puts these two agents on the same A-curve). Indeed, if they were not placed on the same A-curve, then we could remove ϵ of the good from one agent and give it to the other, thereby improving the value of $\mathcal{V}^{\Sigma m}$, contradicting the optimality of X^*. (This uses the fact that whichever agent has a higher value of A must be receiving a positive quantity of the good.) Likewise, if X^* allocates weakly more than $\widehat{X}$ to one of the two agents $(c, e^*), (c, e)$, then it must place these two agents on the same B-curve as each other (and therefore both of them receive weakly more under X^* than under $\widehat{X}$): otherwise, a small transfer would improve the sum of their A-values, and then a further small transfer from either of them to agent (c^*, e^*) would result in a strict improvement in the value of $\mathcal{V}^{m\Sigma}$, contradicting optimality.

Since X^* and $\widehat{X}$ allocate the same total quantity $\overline{x}$, there must be some agent who receives weakly more under X^* than $\widehat{X}$. The above arguments, applied repeatedly, then show that this is the case for all four agents, and so $X^* = \widehat{X}$. But then the same argument also implies that X^* should place agents (c^*, e^*) and (c^*, e) on the same B-curve, otherwise a small transfer will improve the value of $\mathcal{V}^{m\Sigma}$. Since $X^*(c^*, e) = x_3(c, e)$, this requires $X^*(c^*, e^*) = x_4(c, e)$. But we already have $X^*(c^*, e^*) = \widehat{X}(c^*, e^*) = x^* \neq x_4(c, e)$, a contradiction. □

Proof of Proposition 5.7. Suppose that such an advantage aggregator exists; we will identify two mutually contradictory inequalities that it has to satisfy. The construction uses an advantage specification based on the example (1.6) from the introduction, appropriately modified to satisfy Assumption 4.1. All logarithms below refer to natural logarithms.

Let a_1, a_2, b_1, b_2, and δ be numbers chosen to satisfy the conditions (A.20), (A.21) and (A.22) below. (We can verify that, for example, $(a_1, a_2, b_1, b_2, \delta) = (40, 41, 53, 50, 4)$ will work.)

First, require

$$1 < a_1 < a_2 < b_2 < b_1 \tag{A.20}$$

and

$$1 < \frac{a_1}{b_1} \cdot \frac{\log(b_1/b_2)}{\log(a_2/a_1)} < \frac{a_2}{b_2} \cdot \frac{\log(b_1/b_2)}{\log(a_2/a_1)} < 2. \tag{A.21}$$

This allows us to define

$$c_1 = \frac{a_1}{b_1} \cdot \frac{\log(b_1/b_2)}{\log(a_2/a_1)} - 1, \qquad e_1 = \frac{a_1}{b_1}, \qquad c_1' = e_1' = 1,$$

$$c_2 = \frac{a_2}{b_2} \cdot \frac{\log(b_1/b_2)}{\log(a_2/a_1)} - 1, \qquad e_2 = \frac{a_2}{b_2}, \qquad c_2' = e_2' = 1.$$

Note that (A.20–A.21) imply all these expressions are positive, and $c_1 < c_1'$, $e_1 < e_1'$, $c_2 < c_2'$, and $e_2 < e_2'$.

We also require

$$\max\left\{\frac{1}{c_1 e_1}, \log(b_1), \frac{1}{c_2 e_2}, \log(b_2)\right\} < \delta < \min\left\{\frac{\log(b_2)}{c_1}, \frac{\log(a_1)}{e_1}, \frac{\log(b_2)}{c_2}, \frac{\log(a_1)}{e_2}\right\}. \tag{A.22}$$

Now, consider the following specification of advantage. Take $(\underline{c}, \overline{c}) = (\underline{e}, \overline{e}) = (0, 3)$. Let $\psi_1^c : \mathbb{R} \to \mathbb{R}_{++}$ be any strictly increasing, smooth function with

$$\psi_1^c(1) = c_1, \qquad \psi_1^c(2) = c_1', \qquad \psi_1^c(0) \text{ close to } c_1,$$

and $\psi_1^e : \mathbb{R} \to \mathbb{R}_{++}$ be any strictly decreasing, smooth function with

$$\psi_1^e(1) = e_1', \qquad \psi_1^e(2) = e_1, \qquad \psi_1^e(3) \text{ close to } e_1,$$

such that $\psi_1^c(0) \cdot \psi_1^e(3) > 1/\delta$ (possible by (A.22)). Define

$$A_1(c, e, x) = 1 - \exp(-\psi_1^c(c)\psi_1^e(e)(x + \delta)).$$

We can check that this satisfies Assumption 4.1. (The decreasing function ψ_1^e, and the shift of x by $\delta > 1/\psi_1^c(c)\psi_1^e(e)$, ensure the needed monotonicity of $\partial A_1/\partial x$ with respect to e.) It also has an AB-foliation: as in the introduction, the function

$$U_1(c,e,x) = \frac{1}{\psi_1^e(e)(1-A_1(c,e,x))} = \frac{1}{\psi_1^e(e)\cdot\exp(-\psi_1^c(c)\psi_1^e(e)(x+\delta))}$$

is constant on A-curves and B-curves, thus defines an AB-foliation.

Put

$$\overline{x}_1 = \frac{\log(a_1)}{c_1e_1} + \frac{\log(a_1)}{c_1'e_1} + \frac{\log(b_1)}{c_1e_1'} - 3\delta = \frac{\log(a_2)}{c_1e_1} + \frac{\log(a_2)}{c_1'e_1} + \frac{\log(b_2)}{c_1e_1'} - 3\delta,$$

where the equality of the two expressions follows from the definitions of c_1, e_1, c_1', e_1'. Consider the population $\{1,2\}\times\{1,2\}$ and the two allocations X_1, X_1' given by

$$X_1(1,1) = \frac{\log(b_1)}{c_1e_1'} - \delta, \quad X_1(1,2) = \frac{\log(a_1)}{c_1e_1} - \delta, \quad X_1(2,1) = 0, \quad X_1(2,2) = \frac{\log(a_1)}{c_1'e_1} - \delta$$

and

$$X_1'(1,1) = \frac{\log(b_2)}{c_1e_1'} - \delta, \quad X_1'(1,2) = \frac{\log(a_2)}{c_1e_1} - \delta, \quad X_1'(2,1) = 0, \quad X_1'(2,2) = \frac{\log(a_2)}{c_1'e_1} - \delta.$$

The quantity given to each agent is nonnegative by (A.22) and (A.20), and both allocations feature the same total quantity $\overline{x}$. The resulting advantage allocations are

$$X_1 \to \left(1-\overset{1,1}{1}/b_1, 1-\overset{1,2}{1}/a_1, 1-\overset{2,1}{\exp}(-\delta), 1-\overset{2,2}{1}/a_1\right)$$

and

$$X_1' \to \left(1-\overset{1,1}{1}/b_2, 1-\overset{1,2}{1}/a_2, 1-\overset{2,1}{\exp}(-\delta), 1-\overset{2,2}{1}/a_2\right).$$

Refer to these two advantage allocations as Y_1 and Y_2 respectively.

Now, X_1 is the canonical allocation for the given population and quantity $\overline{x}_1$: this follows from the fact that it achieves the same U-values for agents $(1,1)$, $(1,2)$, and $(2,2)$, while achieving this value for agent $(2,1)$ would require a negative amount of the good, so this agent belongs at a corner. (Use the formulas for e_1, e_1', and inequality (A.22), to verify these claims.) Meanwhile, X_1' is a different allocation that is feasible in this same problem. Therefore, if $\mathcal{V}$ selects the canonical allocation as in (5.1), then

$$\mathcal{V}(Y_1) > \mathcal{V}(Y_2). \tag{A.23}$$

Now we repeat the construction with the 1's and 2's swapped. Let $\psi_2^c : \mathbb{R} \to \mathbb{R}_{++}$ be strictly increasing and smooth with

$$\psi_2^c(1) = c_2, \qquad \psi_2^c(2) = c_2', \qquad \psi_2^c(0) \text{ close to } c_2,$$

and $\psi_2^e : \mathbb{R} \to \mathbb{R}_{++}$ be strictly decreasing and smooth with

$$\psi_2^e(1) = e_2', \qquad \psi_2^e(2) = e_2, \qquad \psi_2^e(3) \text{ close to } e_2,$$

such that $\psi_2^c(0) \cdot \psi_2^e(3) > 1/\delta$ (by (A.22)). Now redefine the advantage function as

$$A_2(c, e, x) = 1 - \exp(-\psi_2^c(c)\psi_2^e(e)(x + \delta)).$$

Again, check that it satisfies Assumption 4.1, and it has an AB-foliation defined by the leaf index

$$U_2(c, e, x) = \frac{1}{\psi_2^e(e)(1 - A_2(c, e, x))} = \frac{1}{\psi_2^e(e) \cdot \exp(-\psi_2^c(c)\psi_2^e(e)(x + \delta))}.$$

Consider now the total quantity

$$\overline{x}_2 = \frac{\log(a_2)}{c_2 e_2} + \frac{\log(a_2)}{c_2' e_2} + \frac{\log(b_2)}{c_2 e_2'} - 3\delta = \frac{\log(a_1)}{c_2 e_2} + \frac{\log(a_1)}{c_2' e_2} + \frac{\log(b_1)}{c_2 e_2'} - 3\delta$$

where, again, the equality follows from the definitions of c_2, e_2, c_2', e_2'. Consider the population $\{1, 2\} \times \{1, 2\}$ and total quantity $\overline{x}_2$, and the two allocations X_2, X_2', where

$$X_2(1,1) = \frac{\log(b_2)}{c_2 e_2'} - \delta, \quad X_2(1,2) = \frac{\log(a_2)}{c_2 e_2} - \delta, \quad X_2(2,1) = 0, \quad X_2(2,2) = \frac{\log(a_2)}{c_2' e_2} - \delta$$

and

$$X_2'(1,1) = \frac{\log(b_1)}{c_2 e_2'} - \delta, \quad X_2'(1,2) = \frac{\log(a_1)}{c_2 e_2} - \delta, \quad X_2'(2,1) = 0, \quad X_2'(2,2) = \frac{\log(a_1)}{c_2' e_2} - \delta.$$

Again, both allocations involve total quantity $\overline{x}_2$, and all individual quantities are non-negative. Check that X_2 leads to advantage allocation Y_2, and X_2' leads to advantage allocation Y_1.

Now, similarly to before, we verify X_2 is the canonical allocation for the given population and quantity $\overline{x}_2$, while X_2' is a different allocation with the same total quantity.

Consequently, we must have

$$\mathcal{V}(Y_2) > \mathcal{V}(Y_1). \tag{A.24}$$

Evidently, (A.23) and (A.24) are incompatible. $\square$

References

Ales, Laurence and Christopher Sleet (2022) "Optimal Taxation of Income-Generating Choice," *Econometrica*, **90** (5), 2397–2436.

Debreu, Gerard (1959) "Topological methods in cardinal utility theory," in Arrow, Kenneth J, Samuel Karlin, and Patrick Suppes eds. *Mathematical Methods in the Social Sciences*: Stanford University Press, 16–26.

Diamond, Rebecca (2016) "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000," *American Economic Review*, **106** (3), 479–524.

Dworczak, Piotr ⓡ Scott Duke Kominers ⓡ Mohammad Akbarpour (2021) "Redistribution through markets," *Econometrica*, **89** (4), 1665–1698.

Dworkin, Ronald (1981a) "What is Equality? Part 1: Equality of Welfare," *Philosophy & Public Affairs*, 185–246.

——— (1981b) "What is Equality? Part 2: Equality of Resources," *Philosophy & Public Affairs*, 283–345.

Fleurbaey, Marc (1994) "On fair compensation," *Theory and Decision*, **36**, 277–307.

——— (1995) "Three solutions for the compensation problem," *Journal of Economic Theory*, **65** (2), 505–521.

——— (2008) *Fairness, responsibility, and welfare*: Oxford University Press.

Fleurbaey, Marc and François Maniquet (2011) "Compensation and responsibility," in *Handbook of Social Choice and Welfare*, **2**: North-Holland, 507–604.

——— (2018) "Optimal income taxation theory and principles of fairness," *Journal of Economic Literature*, **56** (3), 1029–1079.

Gorman, William M (1968) "The structure of utility functions," *The Review of Economic Studies*, **35** (4), 367–390.

Kaplow, Louis and Steven Shavell (2001) "Any non-welfarist method of policy assessment violates the Pareto principle," *Journal of Political Economy*, **109** (2), 281–286.

Kranich, Laurence (1996) "Equitable opportunities: an axiomatic approach," *Journal of Economic Theory*, **71** (1), 131–147.

Lockwood, Benjamin B and Matthew Weinzierl (2015) "De Gustibus non est Taxandum: Heterogeneity in preferences and optimal redistribution," *Journal of Public Economics*, **124**, 74–80.

Maniquet, François (2004) "On the equivalence between welfarism and equality of opportunity," *Social Choice and Welfare*, **23** (1), 127–147.

Nigai, Sergey (2016) "On measuring the welfare gains from trade under consumer heterogeneity," *The Economic Journal*, **126** (593), 1193–1237.

Ok, Efe A (1997) "On opportunity inequality measurement," *Journal of Economic Theory*, **77** (2), 300–329.

Ooghe, Erwin, Erik Schokkaert, and Dirk Van de gaer (2007) "Equality of opportunity versus equality of opportunity sets," *Social Choice and Welfare*, **28** (2), 209–230.

Rawls, John (1971) *A Theory of Justice*: Harvard University Press.

Roemer, John E (1993) "A pragmatic theory of responsibility for the egalitarian planner," *Philosophy & Public Affairs*, 146–166.

——— (1998) *Equality of Opportunity*: Harvard University Press.

——— (2002) "Equality of opportunity: A progress report," *Social Choice and Welfare*, 455–471.

Roemer, John E and Alain Trannoy (2016) "Equality of opportunity: Theory and measurement," *Journal of Economic Literature*, **54** (4), 1288–1332.

Van de gaer, Dirk (1993) *Equality of opportunity and investment in human capital* Ph.D. dissertation, KU Leuven.