# DISCUSSION PAPER SERIES

DP14783

## AGGREGATING WELFARE GAINS

Maya Eden

**MACROECONOMICS AND GROWTH**

ISSN 0265-8003

# AGGREGATING WELFARE GAINS

*Maya Eden*

Discussion Paper DP14783
Published 19 May 2020
Submitted 18 May 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

# AGGREGATING WELFARE GAINS


## Abstract

I characterize the set of policy decision rules which, in addition to satisfying the standard Pareto condition and a weak anonymity requirement, utilize only information about the objective effects of a policy change and its associated profile of individual "welfare gains". I establish that, depending on the assumptions made about individual preferences, there is either no social decision rule that satisfies these requirements, or a unique one. I characterize the unique social decision rules under different common definitions of "welfare gains" and common assumptions about individual preferences.



JEL Classification: N/A

Keywords: N/A



Maya Eden - meden@brandeis.edu
*Brandeis University and CEPR*



Acknowledgements
I am grateful for helpful conversations with Uzi Segal and John Weymark, and for comments from seminar participants at the University of Oklahoma.

In his Presidential Address to the American Economic Association, Robert Lucas Jr. presents the following definition of the *welfare gain* from a policy change:[1]

> Suppose we want to compare the effects of two policies, $A$ and $B$ say, on a single consumer. Under policy $A$ the consumer's welfare is $U(c_A)$, where $c_A$ is the consumption level he enjoys under that policy, and under policy $B$ it is $U(c_B)$. Suppose that he prefers $c_B$: $U(c_A) < U(c_B)$. Let $\lambda > 0$ solve
>
> $$U(c_B) = U((1+\lambda)c_A)$$
>
> We call this number $\lambda$ – in units of a percentage of all consumption goods – the *welfare gain* of a change in policy from $A$ to $B$.

This definition of welfare gains, which follows a long tradition in quantitative public finance, is ubiquitous in the macroeconomics literature (Lucas continues by computing the potential welfare gains from eliminating business cycle fluctuations). Other, similar approaches define "welfare gains" based on proportional increases in single good (as in Jones and Klenow [2016]), or as nominal equivalent or compensating variations. Generally, it is common practice in economic policy evaluation to associate a policy change with a profile of individual welfare gains, $(\lambda_1, ..., \lambda_I)$, where $\lambda_i$ is the welfare gain to individual $i$. Social welfare analysis is then performed by aggregating the individual welfare gains in some manner.

This paper sets out to clarify the properties of social choice based on this procedure. I introduce the Independence of Irrelevant Variations (IIV) axiom, which states that the social decision rule must depend only on the profile of individual welfare gains – and not on other aspects of individual preferences. Loosely speaking, the IIV axiom is similar to Arrow's Independence of Irrelevant Alternatives (IIA) axiom (Arrow [1950]), only that it permits welfare comparisons to depend on the strength of individual preferences as captured by some measure of individual welfare gains, rather than just on individuals' binary preference rankings.

The main result of this paper is that the IIV axiom is quite restrictive: depending on the domain of individual preferences, there may be no social preference relation that jointly satisfies the IIV axiom, the standard Pareto condition and a weak anonymity condition. Furthermore, whenever such a social preference relation exists,

[1] See Lucas [2003].

it is unique, and can be represented by a social welfare function that is equal to a sum of appropriately-normalized individual utility functions.

This uniqueness result allows for sharp characterizations of social preference relations under various common assumptions about individual preferences. For example, under the assumption that individual preferences are homothetic, there exists a unique social preference relation that is consistent with the Pareto principle, the weak anonymity requirement, and the assumption that the profile of individual welfare gains as defined in Lucas [2003] is a sufficient statistic for making policy decisions. This unique decision rule implies that policy $B$ is more desirable than policy $A$ if and only if $\sum_{i=1}^{I} \ln(1+\lambda_i) > 0$, where $\lambda_i$ is the welfare gain of individual $i = 1, ..., I$ from switching from $A$ to $B$.

This paper is related to the literature on Arrow's impossibility theorem in economic environments. Arrow's impossibility theorem states that there is no efficient and fair social decision rule that uses only individuals' binary preference rankings. Bailey [1979], Donaldson and Roemer [1987], Donaldson and Weymark [1988], Campbell [1992] and others point out that, as Arrow assumes that any preference ranking is possible, his result is not immediately applicable to economic settings in which the preference domain is restricted to include only increasing, continuous preferences. However, they establish that variants of Arrow's result hold even under these stronger domain restrictions (see Le Breton and Weymark [2011] for a review). Fleurbaey et al. [2005] further extend this impossibility result to settings in which the social decision rule is allowed to depend on local properties of individual preferences (such as marginal rates of substitution between goods). This paper contributes to this literature by establishing a variant of Arrow's impossibility result in settings in which the social decision rule is allowed to depend on individual "welfare gains" – a non-local property of individual preferences which is typically used by economists to convey information about preference "intensity".

There is also a branch of the literature exploring how relaxing Arrow's axioms can lead to possibility results in economic environments (see, among others, Osborne [1976], Fleurbaey and Maniquet [2011] and Brandl and Brandt [2020]). In this paper, I derive a social preference relation based on a relaxation of Arrow's Independence of Irrelevant Alternatives axiom and additional restrictions on the individual preference domain.

Finally, this paper contributes to the literature on the possibility of social choice based on ordinal properties of individual preferences. Harsanyi [1955], Osborne [1976] and others consider social decision rules that depend on interpersonally-comparable quantities of "utility" (such as utilitarianism). However, Malmquist [1953], Kannai [1970], Samuelson [1977], Pazner and Schmeidler [1978], Pazner [1979] and Fleurbaey and Maniquet [2017] illustrate that a cardinal interpretation of individual well-being is not strictly necessary; rather, an interpersonally-comparable measure of utility can be obtained by mapping each individual allocation into a welfare-equivalent allocation in some index.[2] The question then becomes how to choose this index (see, for example, Appendix A in Fleurbaey and Blanchet [2013]). This paper contributes to this literature by offering an alternative approach based on individual welfare gains, which circumvents the problem of index choice and corresponds more closely to the type of information typically reported in economic policy evaluation.

# 1 Preliminaries

There are $2 \leq I < \infty$ individuals indexed $i = 1, ..., I$. The set of individual allocations is $X = \mathbb{R}_+^J$, where $2 \leq J \leq \infty$ is the number of goods. Elements of $X$ will be denoted by $\vec{x} = (x_1, ..., x_J) \in X$. Individuals have self-regarding preferences over elements of $X$, denoted by $\preceq_P$, where $P$ is an index of the individual preference relation.

Throughout, I use bold letters to denote vectors of length $I$ (the number of individuals). A social state $\mathbf{x} = (\vec{x}_1, ..., \vec{x}_I) \in X^I$ represents a state in which the bundle allocated to individual $i$ is $\vec{x}_i = (x_{i,1}, ..., x_{i,J}) \in X$. A preference profile in which the preferences of individual $i$ are given by $P_i$ is denoted $\mathbf{P} = (P_1, ..., P_I)$.

Let $D^*$ denote the set of all continuous and increasing individual preferences. Let $D \subseteq D^*$ denote the individual preference domain, which is the set of individual preferences that are considered reasonable or likely. I restrict attention to preference domains that consist of at least two distinct preferences ($|D| > 1$).

A *constitution* (also known as an Arrow social welfare function) is a set of social rankings of $X^I$, $\{\preceq_{\mathbf{P}}\}_{\mathbf{P} \in D^I}$, where the social preference ranking $\preceq_{\mathbf{P}}$ corresponds to the social preference relation given the profile of individual preferences $\mathbf{P} \in D^I$.[3] A *social welfare function* is a correspondence $W : X^I \times D^I \mapsto \mathbb{R}$ such that $W(\cdot, \mathbf{P})$

[2]This approach is also referred to as the "ray utility approach".

[3]This definition of a constitution assumes that the social ranking of two alternatives is independent of the feasible set of alternatives; this property was formulated as the "Independence of Feasible Set" axiom in Fleurbaey and Maniquet [2011].

is a Bergson-Samuelson social welfare function that represents the social preference relation $\preceq_{\mathbf{P}}$.

Note that this construction assumes that social preferences may depend only on the ordinal properties of the individual preference rankings, $(\preceq_{P_1}, ..., \preceq_{P_I})$. A more welfarist approach would argue that, even if two individuals have the same ordinal preferences, they may have different capabilities of deriving utility from the same bundle of goods, and that this may matter for social decision-making.[4] Another view, advocated by Rawls [1971] and Dworkin [1981], contends that individuals should be held responsible for their ability to derive utility out of the resources allocated to them, and that, consequently, only the ordinal properties of individual preferences should matter for social choice. This latter view is reflected in the construction here.

## 2 Welfare Gains

The literature features various definitions of individual welfare gains. While Lucas [2003] defines welfare gains based on welfare-equivalent proportional increases in all goods, others, such as Jones and Klenow [2016], define welfare gains based on welfare-equivalent proportional increases in a single good (holding all other goods constant). Another approach, which is common in microeconomics, is to report nominal equivalent or compensating variations, and interpret those as representing individual welfare gains. This section offers a general definition of welfare gains which includes all of these approaches as special cases.

Common to all of these approaches is the assumption that, within restricted sets of allocations, "welfare gains" are independent of individual preferences. For example, in Lucas [2003], the welfare gain from increasing all goods by a factor of $(1+\lambda)$ is simply $\lambda$ – regardless of the individual preference relation. Similarly, the equivalent variation of increasing income by $\Delta m$ without changing prices is always $\Delta m$, irrespective of the individual preference relation.

To formalize this property, it is useful to introduce the following notation. For a function $\lambda : X \times X \times D^* \mapsto \mathbb{R}$ and a bundle $\vec{x} \in X$, define

$$S(\vec{x}) = \{\vec{x'} | \lambda(\vec{x}, \vec{x'}, P) = \lambda(\vec{x}, \vec{x'}, P') \text{ and } \lambda(\vec{x'}, \vec{x}, P) = \lambda(\vec{x'}, \vec{x}, P') \ \forall P, P' \in D^*\}$$

[4] For example, a utilitarian approach would allocate more goods to individuals who are capable of deriving more utility out of them (see Harsanyi [1955] and Moulin and Thomson [1997] for a discussion).

The set $S(\vec{x})$ comprises of all bundles $\vec{x'} \in X$ for which the values of $\lambda(\vec{x'}, \vec{x}, P)$ and $\lambda(\vec{x}, \vec{x'}, P)$ do not depend on $P$. For example, if $\lambda$ is the measure of welfare gains as defined in Lucas [2003], then $S(\vec{x})$ consists of all bundles of goods that are proportional to the bundle of goods $\vec{x}$.

In addition, let $\iota(\vec{x})$ be given by

$$\iota(\vec{x}) = \cap_{\vec{x'} \in S(\vec{x})} S(\vec{x'}) \tag{1}$$

The set $\iota(\vec{x})$ is defined so that, for every $\vec{x'}, \vec{x''} \in \iota(\vec{x})$, the value of $\lambda(\vec{x'}, \vec{x''}, P)$ is independent of $P$. For example, if $\lambda$ is the measure of welfare gains as defined in Lucas [2003], then $\iota(\vec{x})$ includes all bundles of goods that proportional to $\vec{x}$, because the welfare gains from switching between any two proportional bundles are independent of the individual preference relation.

I propose the following definition of welfare gains.

**Definition.** *A function $\lambda : X \times X \times D^* \mapsto \mathbb{R}$ is said to be a measure of welfare gains if the following conditions hold:*

1. *For each $\vec{x} \in X$ and $P \in D^*$, the functions $\lambda(\vec{x}, \cdot, P)$ and $-\lambda(\cdot, \vec{x}, P)$ are representations of the preferences $P$.*
2. *For each $\vec{x} \in X$ and $P, P' \in D^*$ it holds that $\lambda(\vec{x}, \vec{x}, P) = \lambda(\vec{x}, \vec{x}, P')$.*
3. *For each $\vec{x} \in X$ and $P \in D^*$, it holds that $\lambda(\vec{x}, \iota(\vec{x}), P) = \lambda(X, X, P)$.*

The first part of the definition requires welfare gains to be monotone in individual preferences. The welfare gain from switching between $\vec{x}$ and $\vec{x'}$ is increasing in the subjective desirability of $\vec{x'}$, and decreasing in the subjective desirability of $\vec{x}$. The second part of the definition is a normalization, requiring that the welfare gains from staying with the same bundle are independent of preferences.

The third part of the definition requires that any possible value of welfare gains can be obtained by switching from any given bundle $\vec{x}$ to a bundle in the set $\iota(\vec{x})$. The following lemma uses this property to establish that, if $\lambda$ is a measure of welfare gains, then each $\iota(\vec{x})$ is a strictly increasing, unbounded index in $\mathbb{R}^J_+$.

**Lemma 1.** *If $\lambda$ is a measure of welfare gains, then*

1. *The sets $\{\iota(\vec{x})\}_{\vec{x} \in X}$ constitute a partition of $X$ into weakly increasing, continuous indexes.*

2. *There exists a unique mapping $e : X \times X \times D^* \mapsto X$ with the property that, for each $\vec{x}, \vec{x'} \in X$ and $P \in D^*$, it holds that $e(\vec{x}, \vec{x'}, P) \sim_P \vec{x'}$ and $e(\vec{x}, \vec{x'}, P) \in \iota(\vec{x})$.*

The proof, together with other omitted proofs, is detailed in the appendix. By definition, $e(\vec{x}, \vec{x'}, P)$ is a bundle in the index $\iota(\vec{x})$ which is welfare-equivalent to $\vec{x'}$ for an individual with preferences $P$. I therefore refer to $e$ as the *welfare-equivalence* relation. Figure 1.A presents an illustration of this construction: axes represent two goods, 1 and 2. The sets $\iota(\vec{x})$ and $\iota(\vec{x'})$ are increasing, unbounded curves, which satisfy $\vec{x} \in \iota(\vec{x})$ and $\vec{x'} \in \iota(\vec{x'})$. In the figure, $IC(P)$ is the indifference curve implied by the preferences $P$, which goes through the bundle $\vec{x'}$. Its intersection with the index $\iota(\vec{x})$ determines the bundle $e(\vec{x}, \vec{x'}, P)$.

Figure 1.B illustrates the partition into indexes implied by the definition of welfare gains in Lucas [2003]. In Lucas, the welfare gains from switching from a consumption bundle $\vec{x}$ to a consumption bundle $\vec{x'}$ is defined based on the indifference condition

$$\vec{x'} \sim_P (1 + \lambda(\vec{x}, \vec{x'}, P))\vec{x} \tag{2}$$

The value of $(1 + \lambda(\vec{x}, \vec{x'}, P))$ is the proportional increase in the consumption bundle $\vec{x}$ that leaves an individual with preferences $P$ indifferent with respect to switching to the consumption bundle $\vec{x'}$.

**Proposition 1.** *The function $\lambda$ defined by expression 2 is a measure of welfare gains. The partition into indexes is given by $\iota(\vec{x}) = \{a\vec{x} | a \in \mathbb{R}_+\}$, and the welfare-equivalence relation is given by*

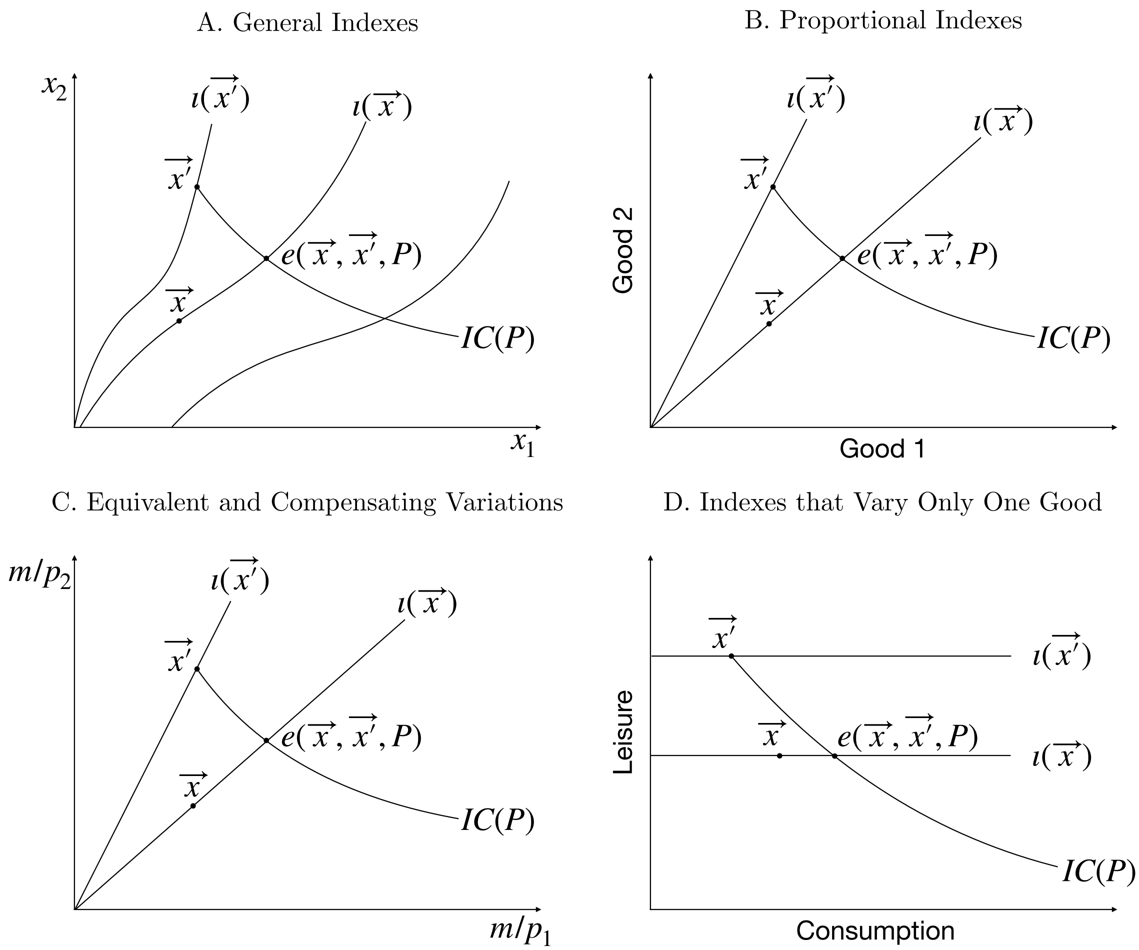
$$e(\vec{x}, \vec{x'}, P) = (1 + \lambda(\vec{x}, \vec{x'}, P))\vec{x} \tag{3}$$

Figure 1.C illustrates the partition into indexes implied by the definition of welfare gains as equivalent variations.[5] Note that, as individuals derive utility out of the consumption of goods rather than directly out of income and prices, their indirect preferences over income ($m$) and prices ($\vec{p} = (p_1, ..., p_J)$) imply unique preferences over vectors of purchasing power of the form $(m/p_1, ..., m/p_J)$. Define the function $\lambda$ as

$$\lambda((\frac{m}{p_1}, ..., \frac{m}{p_J}), (\frac{m'}{p'_1}, ..., \frac{m'}{p'_J}), P) = ev/m \tag{4}$$

[5]Note that the compensating variation is defined the same way as the equivalent variation, with the relabelling of the initial state as the final state and vice-versa. The example of the equivalent variation therefore applies also to the compensating variation, subject to this relabelling.

Figure 1: Indexes and welfare equivalence relations



where $ev$ is the equivalent variation associated with a policy that changes income from $m$ to $m'$ and prices from $(p_1, ..., p_J)$ to $(p'_1, ..., p'_J)$ (given individual preferences $P$). Note that the combination of $\lambda$ and the initial income level $m$ is sufficient for recovering $ev$. Thus, for a given initial income level, the information contained in the equivalent variation ($ev$) is the same as the information contained in the value of $\lambda$.

**Proposition 2.** *Let $\lambda$ be given by equation 4. Then, $\lambda$ is a measure of welfare gains, which implies the partition $\iota(\frac{m}{p_1}, ..., \frac{m}{p_J}) = \{(\frac{m'}{p_1}, ..., \frac{m'}{p_J}) | m' \in \mathbb{R}_+\}$.*

This proposition establishes that the partition into indexes implied by the definition of welfare gains as equivalent variations is similar to the partition implied by the definition of welfare gains in Lucas [2003]. The two examples differ only in the interpretation of $X$: in Lucas [2003], a vector $\vec{x} \in X$ represents the vector of consumption goods, whereas, here, a vector $\vec{x} \in X$ represents a vector of purchasing powers.

The final example, in Figure 1.D, corresponds to the definition of welfare gains used in Jones and Klenow [2016]. Jones and Klenow [2016] compare the economic well-being in countries that vary both in their consumption levels, $x_{i,1} \in \mathbb{R}_+$, and in other economic indicators (such as leisure and life expectancy), $x_{i,j} \in \mathbb{R}_+$ for $j = 2, ..., J$. The welfare of country $i$ is reported as $\lambda_i$, where $\lambda_i$ is defined so that individuals (who are assumed to have common preferences) are indifferent between the bundle $(x_{i,1}, ..., x_{i,J})$ and the bundle $(\lambda_i x_{us,1}, x_{us,2}, ..., x_{us,J})$, where $x_{us,1}, ..., x_{us,J}$ are the values of the different economic indicators in the United States.

Define the function $\lambda(\vec{x}, \vec{x}', P)$ based on the indifference condition

$$\vec{x'} \sim_P (\lambda(\vec{x}, \vec{x'}, P) x_1, x_2, ..., x_J) \tag{5}$$

Note that this definition corresponds to the definition of welfare gains in Jones and Klenow [2016], which uses $\vec{x} = \vec{x}_{us}$ as a benchmark.

**Proposition 3.** *If $\lambda$ satisfies the condition in expression 5, then $\lambda$ is a measure of welfare gains. The partition into indexes is given by $\iota(\vec{x}) = \{(ax_1, x_2, ..., x_J) | a \in \mathbb{R}_+\}$, and the welfare-equivalence relation is given by $e(\vec{x}, \vec{x'}, P) = (\lambda(\vec{x}, \vec{x'}, P) x_1, x_2, ..., x_J)$.*

These examples illustrate that there are many plausible definitions of "welfare gains". This raises the obvious question of what is the *right* notion of welfare gains? While this question is highly relevant for the characterization of the desirable social welfare function, it is beyond the scope of this paper. Rather, this paper starts from a given measure of welfare gains and asks, under what conditions, does this measure provide sufficient information for social decision making.

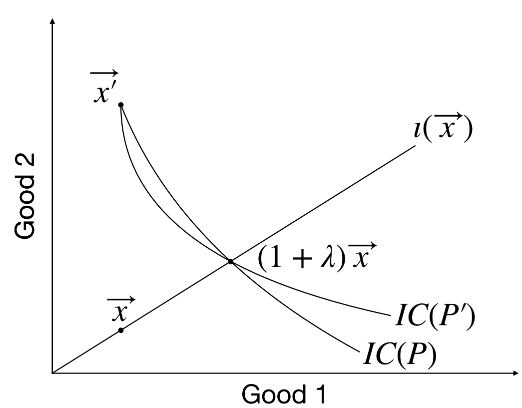
# 3 Axioms

In what follows, I introduce three axioms on the constitution: Independence of Irrelevant Variations, the Pareto condition and Weak Anonymity.

**Independence of Irrelevant Variations.** The Independence of Irrelevant Variations axiom states that all of the information that is relevant for policy decisions is contained in the profile of individual welfare gains, as well as the initial and final states $\mathbf{x}$ and $\mathbf{x}'$:

**Definition.** *Given a measure of welfare gains, $\lambda$, a constitution is said to satisfy the Independence of Irrelevant Variations (IIV) axiom if, for every $\mathbf{x}, \mathbf{x}' \in X^I$ and $\mathbf{P}, \mathbf{P}' \in D^I$ such that, for each $i$, $\lambda(\vec{x}_i, \vec{x'}_i, P_i) = \lambda(\vec{x}_i, \vec{x'}_i, P'_i)$, it holds that $\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}'$ if and only if $\mathbf{x} \preceq_{\mathbf{P}'} \mathbf{x}'$.*

Figure 2: Independence of Irrelevant Variations




The IIV axiom states that the profile of individual welfare gains is a sufficient statistic for conducting welfare analysis. Figure 2 illustrates this graphically. The curves labelled $IC(P)$ and $IC(P')$ represent indifference curves through the consumption bundle $\vec{x'}$, given the individual preferences $P$ and $P'$, respectively. The two indifference curves have the same intersection with the set $\iota(\vec{x})$, which, in this case, corresponds to a ray (as in the Lucas [2003] example): both preferences imply indifference between $\vec{x'}$ and $(1+\lambda)\vec{x} = e(\vec{x}, \vec{x'}, P) = e(\vec{x}, \vec{x'}, P')$. The IIV axiom states that the desirability of a policy which changes an individual's consumption bundle from $\vec{x}$ to $\vec{x'}$ shouldn't depend on whether his preferences are $P$ or $P'$: for example, for each $(\vec{x}_2, ..., \vec{x}_I), (\vec{x'_2}, ..., \vec{x'}_I) \in X^{I-1}$, if $(\vec{x}, \vec{x}_2, ..., \vec{x}_I) \prec_{(P,...,P)} (\vec{x'}, \vec{x'}_2, ..., \vec{x'}_I)$, then $(\vec{x}, \vec{x}_2, ..., \vec{x}_I) \prec_{(P',P,...,P)} (\vec{x'}, \vec{x'}_2, ..., \vec{x'}_I)$.

There are two alternative motivations for the IIV axiom. The first is as an ethical requirement: the only information that *should* matter for assessing the desirability of a policy change is (a) how the objective circumstances of individuals change (the change from $\mathbf{x}$ to $\mathbf{x}'$) and (b) the individual welfare gains associated with this change. For example, it is reasonable to postulate that the decision of whether or not to sign a new trade agreement should depend only on how it will affect prices and income levels, and the individual "welfare gains" from these changes. Other information about individual preferences should have no impact on the decision.

It is worth highlighting that the IIV axiom is sufficiently flexible to allow for social decision rules that value individual welfare gains differentially depending on an individual's objective circumstances. For example, society may place higher value on welfare gains to poorer individuals (those with lower $\vec{x}_i$ or $\vec{x'}_i$).[6] The IIV axiom merely states that the social desirability of implementing a policy shouldn't depend on other aspects of individual preferences, such as the welfare gains from implementing a third policy alternative which is not under consideration.

The second motivation for the IIV axiom is as an informational constraint: while policy makers may like to make decisions based on richer information about individual preferences, the only information available to them is information about individual welfare gains, as reported to them by economists. Thus, for pragmatic reasons, the social decision rule must satisfy the IIV axiom.

**Pareto.** A common requirement of social preferences is that they coincide with individual preferences whenever there is no conflict of interests among individuals: if all individuals agree that a policy is good, then it should be deemed socially desirable. The following standard definition of the Pareto principle (sometimes referred to as "unanimity" or "efficiency") captures this requirement.

[6]To illustrate, note that any constitution defined by a condition of the form $\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}' \Leftrightarrow \sum_{i=1}^{I} \psi(\mathbf{x}, \mathbf{x}', i) f(\lambda(\vec{x}_i, \vec{x'}_i, P_i)) \geq 0$ satisfies the IIV axiom.

**Definition.** *A social preference relation* $\preceq_{\mathbf{P}}$ *is said to be consistent with the Pareto principle (or Paretian) if, for every two social states* $\mathbf{x}, \mathbf{x}' \in X^I$*, it holds that (a) if* $\vec{x}_i \preceq_{P_i} \vec{x'}_i$ *for every* $i$*, then* $\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}'$ *and (b) if, in addition, for some* $i$*,* $\vec{x}_i \prec_{P_i} \vec{x'}_i$*, then* $\mathbf{x} \prec_{\mathbf{P}} \mathbf{x}'$*. A constitution* $\{\preceq_{\mathbf{P}}\}_{\mathbf{P} \in D^I}$ *is said to be consistent with the Pareto principle (or Paretian) if, for every preference profile* $\mathbf{P} \in D^I$*, the social preference relation* $\preceq_{\mathbf{P}}$ *is Paretian.*

**Weak Anonymity.** In addition to the Pareto principle, it is desirable to assume that social preferences are "fair". Intuitively, fairness requires that social preferences are not systematically biased against some individuals, and that the welfare of all individuals is valued equally.

**Definition.** *A constitution is said to satisfy Weak Anonymity if, for every* $1 \leq i < i' \leq I$ *and* $\mathbf{P} \in D^I$ *such that* $P_i = P_{i'}$*,*

$$(\vec{x}_1, ..., \vec{x}_i, ..., \vec{x}_{i'}, ..., \vec{x}_I) \sim_{\mathbf{P}} (\vec{x}_1, ..., \vec{x}_{i'}, ..., \vec{x}_i, ..., \vec{x}_I) \tag{6}$$

Weak anonymity requires the equal treatment of individuals who have the same ordinal preference ranking.[7] This definition of anonymity requires society to be indifferent with respect to switching the consumption bundles of any two individuals ($i$ and $i'$), provided that they have the same preferences. This means that the index number of the individual shouldn't matter for social decision making.

# 4 Characterization of Social Preferences

This section lays out the conditions for the existence and uniqueness of a Paretian constitution that satisfies Weak Anonymity and the IIV axiom.

**Theorem 1.** *Let* $\lambda$ *be a measure of welfare gains, and let* $D$ *be the preference domain.*

1. *Assume that there exist continuous functions* $\{\mu_{\iota(\vec{x})} : \iota(\vec{x}) \mapsto \mathbb{R}\}_{\vec{x} \in X}$ *and functions* $\{\gamma(\cdot|P) : \{\iota(\vec{x})\}_{\vec{x} \in X} \mapsto \mathbb{R}\}_{P \in D}$ *such that, for each* $P \in D$*, the function* $u(\cdot|P) = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$ *is a representation of* $P$*.*

[7] This axiom is similar to the "Anonymity among Equals" axiom in Fleurbaey and Maniquet [2011].

*Then, there exists a unique constitution that jointly satisfies the IIV axiom, the Weak Anonymity condition and the Pareto condition. The constitution can be represented by the social welfare function*

$$W(\mathbf{x}, \mathbf{P}) = \sum_{i=1}^{I} u(\vec{x}_i | P_i) \tag{7}$$

2. *Conversely, if there exists a constitution that jointly satisfies the IIV axiom, the Weak Anonymity condition and the Pareto condition, then there exist continuous functions $\{\mu_{\iota(\vec{x})} : \iota(\vec{x}) \mapsto \mathbb{R}\}_{\vec{x} \in X}$ and functions $\{\gamma(\cdot|P) : \{\iota(\vec{x})\}_{\vec{x} \in X} \mapsto \mathbb{R}\}_{P \in D}$ such that, for each $P \in D$, the function $u(\cdot|P) = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$ is a representation of $P$.*

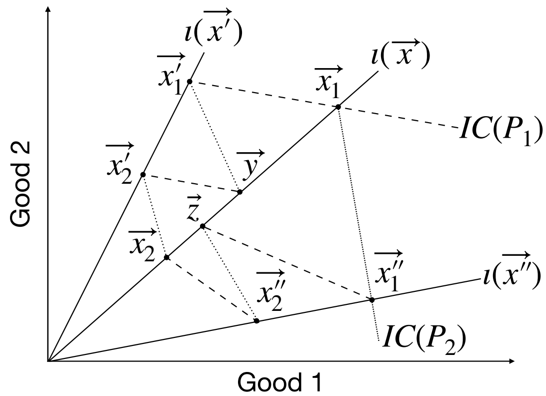
The theorem establishes a necessary and sufficient condition for the existence of a constitution that jointly satisfies the IIV axiom, the Weak Anonymity condition and the Pareto condition. The condition is that each preference relation $P \in D$ can be represented in an additively separable form, $u(\cdot|P) = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$. The first component, $\mu_{\iota(\vec{x})}(\vec{x})$, parameterizes the "utility" of bundles in the index $\iota(\vec{x})$, and the second component, $\gamma(\iota(\vec{x})|P)$, corresponds to an individual ranking of indexes (given some fixed value of $\mu_{\iota(\vec{x})}(\vec{x})$). Note that the first component is required to be the same for all preferences in $D$, while the second component is allowed to depend on the individual preference ranking.[8] At first glance, this separability condition may seem obscure; however, as I establish in the following section, it turns out to correspond to familiar assumptions about individual preferences given the measures of welfare gains discussed in section 2.

The first part of the theorem further establishes that, whenever there exists a constitution that satisfies the theorem's axioms, then it is unique. This result suggests that the theorem's axioms are highly restrictive: depending on the preference domain, there is either no constitution that satisfies them, or only one.

Figure 3 develops intuition and sketches the key steps of the proof. For simplicity, assume that $I = J = 2$, as in Figure 3. In the figure, the curves labelled $IC(P_1)$ and $IC(P_2)$ represent the indifference curves of some preferences $P_1, P_2 \in D$, respectively

[8] Of course, the representations $\{u(\cdot|P)\}_{P \in D}$ are not unique: for example, a common affine transformation $\{\tilde{u}(\cdot|P) = au(\cdot|P) + b\}_{P \in D}$ satisfies the separability conditions above. However, the theorem states that the the resulting social preference relation is the same regardless of the choice of representation, so long as it satisfies the necessary separability condition.

Figure 3: Proof of Thoerem 1




(the dashed curves are indifference curves of $P_1$ and the dotted curves are indifference curves of $P_2$). Note that $\vec{x'}_1 \sim_{P_1} \vec{x}_1$ and $\vec{x_2'} \sim_{P_2} \vec{x}_2$, and thus, by the Pareto principle,

$$(\vec{x}_1, \vec{x}_2) \sim_{(P_1,P_2)} (\vec{x'}_1, \vec{x_2'}) \tag{8}$$

Weak Anonymity requires that, if $P_1 = P_2$, then society should be indifferent with respect to switching the bundles of the two individuals:

$$(\vec{x_2'}, \vec{x'}_1) \sim_{(P_1,P_1)} (\vec{x'}_1, \vec{x_2'}) \tag{9}$$

Note that $\vec{x'}_1, \vec{x_2'}$ were chosen to be in the same index, $\iota(\vec{x'})$. Recall that the defining property of an index is that the welfare gains from switching between any two of its elements do not depend on individual preferences. This implies that $\lambda(\vec{x'}_1, \vec{x_2'}, P_1) = \lambda(\vec{x'}_1, \vec{x_2'}, P_2)$. In addition, the IIV axiom requires that, whenever two preference profiles imply the same profile of individual welfare gains, the social ranking should be the same. This is the case for the two preference profiles $(P_1, P_1)$ and $(P_1, P_2)$: as $\lambda(\vec{x_2'}, \vec{x'}_1, P_1) = \lambda(\vec{x_2'}, \vec{x'}_1, P_1)$ and $\lambda(\vec{x'}_1, \vec{x_2'}, P_1) = \lambda(\vec{x'}_1, \vec{x_2'}, P_2)$, the IIV axiom implies that

$$(\vec{x_2'}, \vec{x'}_1) \sim_{(P_1,P_1)} (\vec{x'}_1, \vec{x_2'}) \Rightarrow (\vec{x_2'}, \vec{x'}_1) \sim_{(P_1,P_2)} (\vec{x'}_1, \vec{x_2'}) \tag{10}$$

Thus, combining with expression 8 yields the indifference condition

$$(\vec{x}_1, \vec{x}_2) \sim_{(P_1,P_2)} (\vec{x_2'}, \vec{x'}_1) \tag{11}$$

Noting that $\vec{x'_2} \sim_{P_1} \vec{y}$ and $\vec{x'}_1 \sim_{P_2} \vec{y}$ and applying the Pareto principle once again yields

$$(\vec{x'_2}, \vec{x'}_1) \sim_{(P_1,P_2)} (\vec{y}, \vec{y}) \tag{12}$$

and, hence, combining with expression 11 yields

$$(\vec{x}_1, \vec{x}_2) \sim_{(P_1,P_2)} (\vec{y}, \vec{y}) \tag{13}$$

This construction illustrates a mapping between an unequal allocation, $(\vec{x}_1, \vec{x}_2) \in \iota(\vec{x})^2$, and an equally-distributed equivalent, $(\vec{y}, \vec{y}) \in \iota(\vec{x})^2$. Note that, as $\vec{x}_1, \vec{x}_2, \vec{y}$ are all elements of the same index $\iota(\vec{x})$, it holds that the profile of welfare gains from switching from $(\vec{x}_1, \vec{x}_2)$ to $(\vec{y}, \vec{y})$ is independent of the profile of individual preference relations, $(P_1, P_2)$. Consequently, the IIV axiom implies that, for every $P'_1, P'_2 \in D$,

$$(\vec{x}_1, \vec{x}_2) \sim_{(P'_1,P'_2)} (\vec{y}, \vec{y}) \tag{14}$$

In other words, the mapping from $(\vec{x}_1, \vec{x}_2)$ to the equally-distributed equivalent, $(\vec{y}, \vec{y})$, cannot depend on the preference profile, $(P_1, P_2)$. It can therefore be written as a function $w : \iota(\vec{x})^2 \mapsto \iota(\vec{x})$, such that

$$(\vec{x}_1, \vec{x}_2) \sim_{(P'_1,P'_2)} (w(\vec{x}_1, \vec{x}_2), w(\vec{x}_1, \vec{x}_2)) \;\forall P'_1, P'_2 \in D \tag{15}$$

Further, note that, by the Pareto principle, for any $(\vec{r}_1, \vec{r}_2) \in X^2$ and $P'_1, P'_2 \in D$, it must hold that $(\vec{r}_1, \vec{r}_2) \sim_{(P'_1,P'_2)} (e(\vec{x}, \vec{r}_1, P'_1), e(\vec{x}, \vec{r}_2, P'_2))$. In other words, society must be indifferent between replacing any arbitrary individual allocation in $X$ with its welfare-equivalent allocation in the index $\iota(\vec{x})$. Given the above expression, it follows that

$$\begin{gathered}(\vec{r_1}, \vec{r_2}) \sim_{(P'_1,P'_2)} (e(\vec{x}, \vec{r}_1, P'_1), e(\vec{x}, \vec{r}_2, P'_2)) \sim_{(P'_1,P'_2)} \\ (w(e(\vec{x}, \vec{r}_1, P'_1), e(\vec{x}, \vec{r}_2, P'_2)), w(e(\vec{x}, \vec{r}_1, P'_1), e(\vec{x}, \vec{r}_2, P'_2)))\end{gathered} \tag{16}$$

As $\iota(\vec{x})$ is a Pareto-ranked set, this establishes a unique social ranking: $(\vec{r}_1, \vec{r}_2) \preceq_{(P'_1,P'_2)} (\vec{s}_1, \vec{s}_2)$ if and only if both individuals weakly prefer the bundle $w(e(\vec{x}, \vec{s}_1, P'_1), e(\vec{x}, \vec{s}_2, P'_2)))$ over the bundle $w(e(\vec{x}, \vec{r}_1, P'_1), e(\vec{x}, \vec{r}_2, P'_2)))$. This complete characterization establishes that, if there exists a constitution that satisfies the axioms of Theorem 1, then it must be unique.

Figure 3 also illustrates why, for certain preference domains, there may be no constitution that satisfies the theorem's axioms: a similar derivation implies that, if the constitution satisfies the theorem's axioms, then $(\vec{x}_1, \vec{x}_2) \sim_{(P_1,P_2)} (\vec{z}, \vec{z})$. However, as $(\vec{z}, \vec{z})$ is Pareto-dominated by $(\vec{y}, \vec{y})$, the combination of the Pareto principle and expression 13 imply that the constitution must satisfy $(\vec{x}_1, \vec{x}_2) \sim_{(P_1,P_2)} (\vec{y}, \vec{y}) \succ_{(P_1,P_2)} (\vec{z}, \vec{z})$. This is a contradiction to the result that $(\vec{x}_1, \vec{x}_2) \sim_{(P_1,P_2)} (\vec{z}, \vec{z})$, establishing that, given this preference profile, there cannot exist a constitution that satisfies the axioms of Theorem 1.

Note that the social welfare function in equation 7 can be interpreted as "utilitarian", in the sense that it is a summation of individual utility functions. However, the theorem requires a particular normalization of individual utility functions which is based only on the ordinal properties of the individual preference relation. The representation does not necessarily correspond to any cardinal notion of the "intensity" of preferences, and is not assumed to represent any interpersonally comparable quantity. In this sense, the social welfare function in equation 7 is not utilitarian, as it is not the sum of individual interpersonally-comparable "utilities".

# 5 Implications

This section presents several corollaries of Theorem 1.

**Corollary 1.** *Let $\lambda$ be a measure of welfare gains, and assume that the preference domain, $D$, is given by the set of increasing, continuous preferences ($D = D^*$). Then, there exists no constitution that jointly satisfies the axioms of Theorem 1.*

This corollary is an impossibility result: there is no constitution that satisfies the axioms of Theorem 1 on the entire domain of increasing and continuous preferences, $D^*$. This implies that further domain restrictions must be imposed in order to make a constitution consistent with the theorem's axioms.

**Corollary 2.** *Let $\lambda$ be the measure of welfare gains defined by expression 2 (as in Lucas [2003]), and let $D$ be a preference domain consisting only of homothetic preferences. Then, there exists a unique constitution that satisfies the axioms of Theorem 1, which can be represented by the social welfare function*

$$W(\mathbf{x}, \mathbf{P}) = \prod_{i=1}^{I} h(\vec{x}_i | P_i) \tag{17}$$

*where $h(\cdot|P) : \mathbb{R}^J_+ \mapsto \mathbb{R}_+$ is any representation of $P$ that is homogeneous of degree 1.*

This corollary establishes that, given a preference domain consisting only of homothetic preferences, the measure of welfare gains in Lucas [2003] implies a unique social preference relation, which can be represented by a product of (any) utility functions that are normalized to be homogeneous of degree 1. Note that the policy recommendation that comes out of the social welfare function in Corollary 2 admits to a simple formula: the policy change is desirable if and only if $\prod_{i=1}^{I}(1+\lambda_i) > 1$, where $\lambda_i$ is the welfare gain to individual $i$.[9]

It is instructive to contrast this social decision rule with the one implied by Harsanyi's theorems (Harsanyi [1953], Harsanyi [1955]). Harsanyi's impartial observer approach remains the most common method for conducting social welfare analysis. To map Harsanyi's setting into the current one, it is convenient to restrict attention to the case $I = J$ (the number of individuals equals the number of "goods"), and interpret the set $X = \mathbb{R}^J_+$ as a set of equiprobable lotteries with $J$ possible outcomes: an element $(x_1, ..., x_J) \in X$ is interpreted as a lottery that assigns probability $1/J$ to each of the consumption levels $x_1, ..., x_J$. When individuals are expected utility maximizers, Harsanyi's social welfare function is given by a weighted sum of individual von-Neumann and Morgenstern utilities. In particular, if, for some preference profile $\mathbf{P} \in D^I$, the preferences $P_i$ are represented by the constant relative risk aversion form $\sum_{i=1}^{J} x_{i,j}^{1-\rho_i}/(1-\rho_i)$, then Harsanyi's social welfare function, $W^H$, is given by

$$W^H(\mathbf{x}, \vec{P}) = \sum_{i=1}^{I} \psi_i \sum_{i=1}^{J} \frac{x_{i,j}^{1-\rho_i}}{1-\rho_i} \tag{18}$$

for some weights $\psi_i \in \mathbb{R}$.

In contrast, the social welfare function implied by Corollary 2 is given by

$$W(\mathbf{x}, \vec{P}) = \prod_{i=1}^{I} (\sum_{i=1}^{J} x_{i,j}^{1-\rho_i})^{\frac{1}{1-\rho_i}} \tag{19}$$

[9] To see this, let $\vec{x}$ be the benchmark allocation and let $\vec{x'}$ be the allocation under the proposed policy change. Note that, by expression 2, $\vec{x'}_i \sim_{P_i} (1+\lambda_i)\vec{x}_i$, and hence $h(\vec{x'}_i|P_i) = h((1+\lambda_i)\vec{x}_i)$. It thus follows that $W(\mathbf{x}, \mathbf{P}) = \prod_{i=1}^{I} h(\vec{x}_i|P_i) \leq \prod_{i=1}^{I} h(\vec{x'}_i|P_i) = W(\mathbf{x}', \mathbf{P})$ if and only if $\prod_{i=1}^{I} h(\vec{x}_i|P_i) \leq \prod_{i=1}^{I} h((1+\lambda_i)\vec{x}_i|P_i) = \prod_{i=1}^{I}(1+\lambda_i) \prod_{i=1}^{I} h(\vec{x}_i|P_i)$, which holds if and only if $1 \leq \prod_{i=1}^{I}(1+\lambda_i)$.

(to see this, note simply that $(\sum_{i=1}^{J} x_{i,j}^{1-\rho_i})^{\frac{1}{1-\rho_i}}$ is a representation of the preferences $P_i$ that is normalized to be homogeneous of degree 1). It is straightforward to see that the two social preference relations are not the same. This is not surprising given that the assumptions used to derive them are very different: in particular, Harsanyi assumes that both individual and social preferences satisfy the expected utility axioms, and that appropriately-normalized von-Neumann Morgenstern utilities represent interpersonally-comparable quantities. None of these assumptions are made here.

Interestingly, the social preference relation represented by equation 19 can be represented in a "generalized utilitarian" form: Grant et al. [2010] show that, by relaxing some of Harsanyi's more controversial assumptions, it is possible to generalize his social welfare function to be of the form

$$W^{GKPS}(\vec{x}, \vec{P}) = \sum_{i=1}^{I} \phi_i(\sum_{i=1}^{J} \frac{x_{i,j}^{1-\rho_i}}{1-\rho_i}) \tag{20}$$

for some functions $\phi_i : \mathbb{R} \mapsto \mathbb{R}$. It is straightforward to show that the social preferences represented by equation 19 can also be represented by this form, by choosing $\phi_i(x) = \ln(((1-\rho_i)x)^{\frac{1}{1-\rho_i}})$.

Finally, it is worth pointing out the relationship between the social preference relation represented by equation 17 and the "Nash social welfare function" derived in Osborne [1976], Kaneko and Nakamura [1979] and Sprumont [2018]. Both social welfare functions are products of individual utility functions; however, they normalize utility functions in different ways. The social welfare function in equation 17 requires individual utility functions to be homogeneous of degree 1, while the Nash social welfare function normalizes individual utility functions by fixing the utility of the worst possible outcome (in a bounded space of alternatives). More closely related, Fleurbaey and Maniquet [2011] (Chapter 6.4) focus on homothetic preference domains and derive the social preference relation represented by equation 17 using an entirely different axiomatic foundation.[10]

The next corollary establishes a procedure for aggregating equivalent variations:

[10] Fleurbaey and Maniquet [2011] do not impose the IIV axiom or the Weak Anonymity axiom; rather, their derivation is based on the Pareto indifference condition, a separability requirement, a continuity requirement and an optimality condition stating that, in distributive problems, it is always optimal to allocate the aggregate endowment equally among individuals.

**Corollary 3.** *Assume that $D$ consists only of individual preferences that are homothetic and convex over the set of bundles of goods, $C = \mathbb{R}^J_+$. Social preferences are defined over the set $X^I$, where $X$ consists of vectors of purchasing powers of the form $(m/p_1, ..., m/p_J)$, where $\vec{p}$ represents a vector of prices and $m$ represents income. Consider a policy that changes the vector of incomes from $(m_1, ..., m_I)$ to $(m'_1, ..., m'_I)$ and the profile of individual prices from $(\vec{p_1}, ..., \vec{p_I})$ to $(\vec{p'_1}, ..., \vec{p'_I})$. Assume that the constitution satisfies the axioms of Theorem 1, and that the measure of welfare gains is given by $\lambda_i = ev_i/m_i$, where $ev_i$ is the individual's equivalent variation (as in equation 4). Then, the policy change is desirable if and only if $\prod_{i=1}^{I}(1 + ev_i/m_i) > 1$.*

This corollary establishes that, if social decisions are to depend only on the profile of equivalent variations (and satisfy the Pareto principle and the Weak Anonymity requirement), then they must follow a simple formula: a policy change is desirable if and only if $\prod_{i=1}^{I}(1 + ev_i/m_i) > 1$, where $ev_i$ is the equivalent variation of individual $i$ and $m_i$ is his initial income.

The social preferences implied by Corollary 3 admit to "Price Independent Welfare Prescriptions" (PIWP): the social ranking of any two income distributions, $(m_1, ..., m_I)$ and $(m'_1, ..., m'_I)$, does not depend on prevailing prices. This property is often implicitly assumed in policy discussions around issues of income inequality and the optimal amount of redistribution – presumably, the optimal amount of redistribution should not depend on the price of rice relative to wheat. To see that the social preference relation in Corollary 3 is consistent with PIWP, note that, holding prices constant, a change in income of $\Delta m_i = m'_i - m_i$ results in an equivalent variation of $ev_i = \Delta m_i$, regardless of the price level. By the above corollary, the change is desirable if and only if $\prod_{i=1}^{I}(1 + \Delta m_i/m_i) > 1$, which is a condition that is independent of prices.

It is worth highlighting that PIWP is a result of the combination of the three axioms of Theorem 1, rather than a hardwired assumption. There is a debate about whether it is appropriate to *require* the social preference relation to satisfy PIWP (see Blackorby et al. [1994] and Fleurbaey and Blanchet [2013], chapter 4). Roberts [1980] finds that requiring the social preference relation to generate price-independent welfare prescriptions imposes strong restrictions on the social welfare function.[11] He

[11] See also Slivinski [1983], who finds that any social welfare criterion that generates price-independent welfare prescriptions when individuals face different prices must take the Cobb-Douglas form.

concludes that "although price-independent welfare prescriptions are widely made, the conditions under which they are theoretically justified are extremely restrictive". Theorem 1 suggests instead that price independent welfare prescriptions are a feature of a social welfare function that is characterized based on three plausible axioms.

The final corollary focuses on the measure of welfare gains in Jones and Klenow [2016]:

**Corollary 4.** *Let $\lambda$ be the measure of welfare gains defined by expression 5, and let $D$ be such that, for some $f_1 : \mathbb{R}_+ \mapsto \mathbb{R}$ and $\{\{f_j(\cdot|P) : \mathbb{R}_+ \mapsto \mathbb{R}\}_{j=2}^J\}_{P \in D}$, each preference relation $P \in D$ can be represented by $f_1(x_1) + \sum_{j=2}^J f_j(x_j|P)$. Then, there exists a unique constitution that satisfies the axioms of Theorem 1, which can be represented by $W((\vec{x}_1, ..., \vec{x}_I), \mathbf{P}) = \sum_{i=1}^I (f_1(x_{i,1}) + \sum_{j=2}^J f_j(x_{i,j}|P_i))$.*

To illustrate, consider an example in which individuals have heterogeneous preferences over consumption ($c$) and leisure ($l$). Specifically, there exist functions $f_c, f_l : \mathbb{R}_+ \mapsto \mathbb{R}$ and preference-specific weights $\{\alpha_P > 0\}_{P \in D}$ such that each $P \in D$ can be represented by the separable utility function

$$u(c, l|P) = f_c(c) + \alpha_P f_l(l) \tag{21}$$

Note that this separability assumption is common in the macro and labor literatures (see, for example, the survey in Keane and Rogerson [2012]).

Consider two alternative measures of individual welfare gains. The first measure defines individual welfare gains as in Jones and Klenow [2016], based on proportional increases in consumption: in expression 5, good 1 is interpreted as consumption and good 2 is interpreted as leisure. Given this measure, the above corollary states that the social preferences must be represented by the social welfare function $\sum_{i=1}^I u(c_i, l_i|P_i)$.

The second measure of individual welfare gains is defined analogously based on proportional increases in leisure rather than consumption: in expression 5, good 1 is interpreted as leisure while good 2 is interpreted as consumption. Note that, given the relabeling of the goods, $u(c, l|P)/\alpha_P = f_l(l) + f_c(c)/\alpha_P$ is a representation of the preferences $P$ which satisfies the conditions of Corollary 4. In this case, the corollary implies a social preference relation that can be represented by the social welfare function $\sum_{i=1}^I u(c_i, l_i|P_i)/\alpha_{P_i}$.

It is straightforward to see that the two social welfare functions represent different social preference relations. In particular, the leisure-based measure of welfare gains implies a social welfare function that puts relatively lower "Pareto weights" on the utilities of individuals who have stronger preferences for leisure (higher $\alpha_P$). Equivalently, the consumption-based measure implies relatively higher "Pareto weights" on individuals who value leisure relatively more. This example illustrates that the choice of how to specify *individual* welfare gains is not innocuous: it may substantively affect the form of the social welfare function.

# 6 Conclusion

Arrow's impossibility theorem establishes that there is no reasonable procedure for choosing between two social alternatives based solely on the profile of the individual rankings of those two alternatives. Intuitively, this is because the binary preference rankings do not contain enough information about *how much* each individual values each of the two alternatives. In economic applications, reported measures of individual welfare gains aim to fill this informational gap.

This paper establishes that, given appropriate domain restrictions, measures of individual welfare gains indeed provide sufficient information for conducting social welfare analysis. Furthermore, the procedure for conducting social welfare analysis based on this information is unique: any other procedure will violate the Pareto condition, the Weak Anonymity condition or the transitivity requirement.

This uniqueness result suggests that an economist's choice of how to report individual welfare gains is not innocuous. One interpretation of the IIV axiom is as an informational constraint: policy makers must make decisions based on the profile of individual welfare gains reported to them by economists, because this is the only information that they have. If policy makers restrict themselves to decision rules that are fair and efficient (in the sense that they satisfy the Weak Anonymity requirement and the Pareto principle), then their decision rule is uniquely pinned down by the type of information that they receive. As different measures of individual welfare gains imply different decision rules, the economist's choice of which measure to report may be highly consequential.

# References

Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950.

Martin J. Bailey. The possibility of rational social choice in an economy. *Journal of Political Economy*, 87(1):37–56, 1979.

Charles Blackorby, François Laisney, and Rolf Schmachtenberg. Ethically consistent welfare prescriptions are reference-price-independent. In I Walker R. Blundell, I. Preston, editor, *The Measurement of Household Welfare*. Cambridge University Press, 1994.

Florian Brandl and Felix Brandt. Arrovian aggregation of convex preferences. *Econometrica*, Forthcoming, 2020.

Donald E. Campbell. Transitive social choice in economic environments. *International Economic Review*, 33(2):341–352, 1992.

David Donaldson and John E. Roemer. Social choice in economic environments with dimensional variation. *Social Choice and Welfare*, 4(4):253–276, 1987.

David Donaldson and John A. Weymark. Social choice in economic environments. *Journal of Economic Theory*, 46(2):291–308, 1988.

Ronald Dworkin. What is equality? part 2: Equality of resources. *Philosophy & Public Affairs*, 10(4):283–345, Autumn 1981.

Marc Fleurbaey and Didier Blanchet. *Beyond GDP: Measuring Welfare and Assessing Sustainability*. Oxford University Press, 2013.

Marc Fleurbaey and François Maniquet. *A theory of fairness and social welfare*. Cambridge University Press, 2011.

Marc Fleurbaey and François Maniquet. Fairness and well-being measurement. *Mathematical Social Sciences*, 90:119–126, 2017.

Marc Fleurbaey, Kotaro Suzumura, and Koichi Tadenuma. Arrovian aggregation in economic environments: How much should we know about indifference surfaces? *Journal of Economic Theory*, 124(1):22–44, 2005.

Simon Grant, Atsushi Kajii, Ben Polak, and Zvi Safra. Generalized utilitarianism and harsanyi's impartial observer theorem. *Econometrica*, 78(6):1939–1971, 2010.

John C. Harsanyi. Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy*, 61(5):434–435, 1953.

John C. Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4):309–321, 1955.

Charles I. Jones and Peter J. Klenow. Beyond gdp? welfare across countries and time. *The American Economic Review*, 106(9):2426–2457, 2016.

Mamoru Kaneko and Kenjiro Nakamura. The nash social welfare function. *Econometrica*, 47(2):423–435, March 1979.

Yakar Kannai. Continuity properties of the core of a market. *Econometrica*, 38(6): 791–815, 1970.

Michael Keane and Richard Rogerson. Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literature*, 50(2): 464–476, March 2012.

Michel Le Breton and John A. Weymark. Arrovian social choice theory on economic domains. *Handbook of social choice and welfare*, 2:191–299, 2011.

Robert E. Lucas. Macroeconomic priorities. *The American Economic Review*, 93(1), 2003.

Sten Malmquist. Index numbers and indifference surfaces. *Trabajos de Estadistica y de Investigacion Operativa*, 4(2):209–242, 1953.

Hervé Moulin and William Thomson. Axiomatic analysis of resource allocation problems. In Sen A. Arrow K.J. and Suzumura K., editors, *Social choice re-examined*, pages 101–120. Palgrave Macmillan, London, 1997.

Dale K. Osborne. Irrelevant alternatives and social welfare. *Econometrica*, 44(5): 1001–1015, September 1976.

Elisha A. Pazner. Equity, nonfeasible alternatives and social choice: A reconsideration of the concept of social welfare. In J.J. Laffont, editor, *Aggregation and Revelation of Preferences*. North-Holland, Amsterdam, 1979.

Elisha A. Pazner and David Schmeidler. Egalitarian equivalent allocations: A new concept of economic equity. *Quarterly Journal of Economics*, 92(4):671–687, November 1978.

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

Kevin Roberts. Price-independent welfare prescriptions. *Journal of Public Economics*, 13(3):277–297, 1980.

Paul A. Samuelson. Reaffirming the existence of 'reasonable' bergson-samuelson social welfare functions. *Economica*, 44:81–88, 1977.

Alan D. Slivinski. Income distribution evaluation and the law of one price. *Journal of Public Economics*, 20(1):103–112, 1983.

Yves Sprumont. Belief-weighted nash aggregation of savage preferences. *Journal of Economic Theory*, 178:222–245, November 2018.

# A Proof of Lemma 1

To prove the first part of the lemma, let $\vec{x} \in X$ and $\vec{x'}, \vec{x''} \in \iota(\vec{x})$. By definition of $\iota(\vec{x})$, it holds that $\vec{x''} \in S(\vec{x'})$ and hence $\lambda(\vec{x'}, \vec{x''}, P) = \lambda(\vec{x'}, \vec{x''}, P')$ for all $P, P' \in D^*$.

Assume by way of contradiction that $x'_j > x''_j$ for some $j$ and $x'_j < x''_j$ for some other $j$. There exist preferences $P, P' \in D^*$ such that $(x'_1, ..., x'_J) \prec_P (x''_1, ..., x''_J)$ and $(x''_1, ..., x''_J) \prec_{P'} (x'_1, ..., x'_J)$; given that $\lambda(\vec{x'}, \cdot, P)$ is a representation of $P$, it follows that $\lambda(\vec{x'}, \vec{x'}, P) < \lambda(\vec{x'}, \vec{x''}, P)$, and, similarly, as $\lambda(\vec{x'}, \cdot, P')$ is a representation of $P'$, it follows that $\lambda(\vec{x'}, \vec{x''}, P') > \lambda(\vec{x'}, \vec{x'}, P')$. Given the normalization requiring that $\lambda(\vec{x'}, \vec{x'}, P) = \lambda(\vec{x'}, \vec{x'}, P')$, it cannot be the case that $\lambda(\vec{x'}, \vec{x''}, P) = \lambda(\vec{x'}, \vec{x''}, P')$, and hence it cannot be the case that $\vec{x'}, \vec{x''} \in \iota(\vec{x})$. This concludes the proof that $\iota(\vec{x})$ is weakly ordered.

To show that $\iota(\vec{x})$ is unbounded, assume by way of contradiction that there exists $a \in \mathbb{R}_+$ such that $(a, ..., a)$ is a strict upper-bound on $\iota(\vec{x})$. Then, as $\lambda(\vec{x}, \cdot, P)$ is a representation of the individual preferences $P$, if holds that $\lambda(\vec{x}, \vec{x'}, P) < \lambda(\vec{x}, (a, ..., a), P)$

for all $\vec{x'} \in \iota(\vec{x})$ – a contradiction to the assumption that $\lambda$ is a measure of welfare gains (third property in the definition) and hence $\lambda(\vec{x}, \iota(\vec{x}), P) = \lambda(X, X, P)$. Similarly, $\iota(\vec{x})$ cannot have a lower bound.

To conclude the proof of the first part of the lemma, it is necessary to show that $\{\iota(\vec{x})\}_{\vec{x}\in X}$ is a partition. To see this, note that $\vec{x} \in \iota(\vec{x})$, and hence for each $\vec{x} \in X$ there exists a set $A \in \{\iota(\vec{x'})\}_{\vec{x'}\in X}$ such that $\vec{x} \in A$.

Thus, to show that $\{\iota(\vec{x})\}_{\vec{x}\in X}$ is a partition, it is left to show that if $\vec{x} \in \iota(\vec{x'}) \cap \iota(\vec{x''})$ then $\iota(\vec{x'}) = \iota(\vec{x''})$. To show this, I establish that if $\vec{x} \in \iota(\vec{x'})$, then $\iota(\vec{x}) = \iota(\vec{x'})$ (and thus, if $\vec{x} \in \iota(\vec{x'}) \cap \iota(\vec{x''})$ then $\iota(\vec{x'}) = \iota(\vec{x}) = \iota(\vec{x''})$).

By definition of $\iota(\vec{x'})$, the assumption that $\vec{x} \in \iota(\vec{x'})$ implies that $\vec{x} \in S(\vec{z})$ for every $\vec{z} \in S(\vec{x'})$. As the condition for inclusion in $S(\vec{z})$ is symmetric, it follows that $\vec{z} \in S(\vec{x})$ for every $\vec{z} \in S(\vec{x'})$, or, equivalently, $S(\vec{x'}) \subseteq S(\vec{x})$. This implies that

$$\iota(\vec{x}) = \cap_{z\in S(\vec{x})} S(z) \subseteq \cap_{z\in S(\vec{x'})} S(z) = \iota(\vec{x'}) \tag{22}$$

To show that $\iota(\vec{x'}) = \iota(\vec{x})$, assume by way of contradiction that there exists $\vec{z} \in \iota(\vec{x'}) \setminus \iota(\vec{x})$. Fix some $P \in D^*$. As $\lambda$ is a measure of welfare gains, there exists $\vec{y} \in \iota(\vec{x})$ such that $\lambda(\vec{z}, \vec{y}, P) = \lambda(\vec{z}, \vec{z}, P)$. As $\lambda(\vec{z}, \cdot, P)$ represents the preferences $P$, this implies that $\vec{z} \sim_P \vec{y}$.

As $\iota(\vec{x})$ is weakly ordered and $\vec{z}, \vec{y} \in \iota(\vec{x'})$, it must hold that either $z_j \leq y_j$ for every $j$ or $y_j \leq z_j$ for every $j$. As $P$ is a strictly increasing preference relation, the finding that $\vec{z} \sim_P \vec{y}$ implies that it must be the case that $z_j = y_j$ for every $j$: otherwise, if, for example, $z_j \leq y_j$ for every $j$ and $\vec{z} \neq \vec{y}$, then, as $P$ is increasing, $\vec{z} \prec_P \vec{y}$. Thus, $\vec{z} = \vec{y} \in \iota(\vec{x})$, in contradiction to the assumption that $\vec{z} \in \iota(\vec{x'}) \setminus \iota(\vec{x})$. It thus follows that $\iota(\vec{x}) = \iota(\vec{x'})$, concluding the proof of the first part of the lemma.

To prove the second part of the lemma, note that the third defining property of a measure of welfare gains guarantees that for each $\vec{x}, \vec{x'} \in X$ and $P \in D^*$ there exists $\vec{x''} \in \iota(\vec{x})$ such that $\lambda(\vec{x}, \vec{x''}, P) = \lambda(\vec{x}, \vec{x'}, P)$. Further, $\vec{x''}$ must be unique: as $\lambda(\vec{x}, \cdot, P)$ is a representation of the preferences $P$, if $\vec{z} \in \iota(\vec{x})$ satisfies $\lambda(\vec{x}, \vec{z}, P) = \lambda(\vec{x}, \vec{x''}, P) = \lambda(\vec{x}, \vec{x'}, P)$, then $\vec{z} \sim_P \vec{x''}$. As $P$ is increasing and $\iota(\vec{x})$ is weakly ordered, if $\vec{z} \sim_P \vec{x''}$ then $\vec{z} = \vec{x''}$. Thus, the mapping $e(\vec{x}, \vec{x'}, P) = \vec{x''}$ is the unique mapping satisfying the desired properties.

## B Proof of Proposition 1

As the first two clauses of the definition of welfare gains are straightforward to verify, I focus on proving that the definition of $\lambda$ in expression 2 satisfies the third clause. This definition implies that the sets $\{\iota(\vec{x})\}_{\vec{x}\in X}$ satisfy the property

$$\{a\vec{x}|a \in \mathbb{R}_+\} \subseteq \iota(\vec{x}) \tag{23}$$

To see this, note that the welfare gains from switching between $a\vec{x}$ and $a'\vec{x}$ are the same for all increasing and continuous preferences relations; it thus follows that, for each $a, a' \in \mathbb{R}$, $a\vec{x} \in S(a'\vec{x})$ and, hence, $\{a\vec{x}|a \in \mathbb{R}\} \subseteq \iota(\vec{x})$.

Note that $\lambda(\vec{x}, \{a\vec{x}|a \in \mathbb{R}\}, P) = \lambda(X, X, P) = (-1, \infty)$. By expression 23, $(-1, \infty) = \lambda(\vec{x}, \{a\vec{x}|a \in \mathbb{R}\}, P) \subset \lambda(\vec{x}, \iota(\vec{x}), P)$ and, as $\lambda(\vec{x}, \iota(\vec{x}), P) \subseteq \lambda(X, X, P) = (-1, \infty)$, it follows that $\lambda(\vec{x}, \iota(\vec{x}), P) = \lambda(X, X, P)$.

The observation that $\{a\vec{x}|a \in \mathbb{R}\} = \iota(\vec{x})$ follows from the first part of Lemma 1, according to which $\iota(\vec{x})$ is a weakly ordered set (note that if $\vec{x'} \neq a\vec{x}$ for every $a \in \mathbb{R}$, there must exist some $a' \in \mathbb{R}$ such that $\vec{x'}$ and $a'\vec{x}$ are not ordered).

Given this partition, the welfare-equivalence relation $e$ is defined as

$$e(\vec{x}, \vec{x'}, P) = (1 + \lambda(\vec{x}, \vec{x'}, P))\vec{x} \tag{24}$$

Note that, by expression 2, it holds that $\vec{x'} \sim_P e(\vec{x}, \vec{x'}, P)$, and, by expression 23, it holds that $e(\vec{x}, \vec{x'}, P) \in \iota(\vec{x})$.

## C Proof of Proposition 2

As the first two clauses of the definition of welfare gains are straightforward to verify, I focus on proving that the definition of $\lambda$ satisfies the third clause. Note that $\lambda$ implies expression 23. To see this, note that, for each $a, a' \in \mathbb{R}$,

$$\lambda(a\vec{x}, a'\vec{x}, P) = \lambda((\frac{am}{p_1}, ..., \frac{am}{p_J}), (\frac{a'm}{p_1}, ..., \frac{a'm}{p_J}), P)$$

where $\vec{x} = (m/p_1, ..., m/p_J)$. Note that the equivalent variation associated with switching from $(\frac{am}{p_1}, ..., \frac{am}{p_J})$ to $(\frac{a'm}{p_1}, ..., \frac{a'm}{p_J})$ is $ev = (a'm - am)$. Thus, $\lambda = ev/m = (a' - a)$, which does not depend on the preference relation $P$. Expression 23 follows.

Similar to the proof of Proposition 1, this property implies that $\lambda$ is a measure of welfare gains and that $\iota(\vec{x}) = \{a\vec{x}|a \in \mathbb{R}_+\}$.

# D Proof of Proposition 3

As the first two clauses of the definition of welfare gains are straightforward to verify, I focus on proving that the definition of $\lambda$ satisfies the third clause. Note that

$$\{(ax_1, x_2, ..., x_J)|a \in \mathbb{R}\} \subseteq \iota(\vec{x}) \tag{25}$$

To see this, let $a, a' \in \mathbb{R}$ and note that $(a'x_1, x_2, ..., x_J) = (a(a'/a)x_1, x_2, ..., x_J)$. It follows that, for each $P \in D^*$,

$$(a'x_1, x_2, ..., x_J) \sim_P (a\frac{a'}{a}x_1, x_2, ..., x_J)$$

This implies that, for each $P \in D^*$, it holds that $\lambda(ax_1, x_2, ..., x_J), (a'x_1, x_2, ..., x_J), P)) = a'/a$, which is independent of $P$.

Thus, for each $\vec{x}$ and $P$, it holds that $\lambda(\vec{x}, \{(ax_1, x_2, ..., x_J)|a \in \mathbb{R}\}, P) = (0, \infty) = \lambda(X, X, P)$. Given that $\lambda$ is a measure of welfare gains, by Lemma 1, the set $\iota(\vec{x})$ must be ordered and hence (following similar steps to the proof of Proposition 1), $\iota(\vec{x}) = \{(ax_1, x_2, ..., x_J)|a \in \mathbb{R}\}$, and $e(\vec{x}, \vec{x'}, P) = (\lambda(\vec{x}, \vec{x'}, P)x_1, x_2, ...., x_J)$.

# E Proof of Theorem 1

**Proof of the first part of the theorem.** To prove existence, note that the constitution represented by equation 7 trivially satisfies the Pareto condition and the Weak Anonymity condition. To establish consistency with the IIV axiom, let there be $\mathbf{x}, \mathbf{x}' \in X^I$ and $\mathbf{P}, \mathbf{P}' \in D^I$ such that, for each $i$, $\lambda(\vec{x_i}, \vec{x'_i}, P_i) = \lambda(\vec{x_i}, \vec{x'_i}, P'_i)$.

As $\lambda(\vec{x_i}, \cdot, P_i)$ is a representation of the preferences $P_i$ and $\vec{x'_i} \sim_{P_i} e(\vec{x_i}, \vec{x'_i}, P_i)$, it follows that $\lambda(\vec{x_i}, \vec{x'_i}, P_i) = \lambda(\vec{x_i}, e(\vec{x_i}, \vec{x'_i}, P_i), P_i)$ and, similarly, $\lambda(\vec{x_i}, \vec{x'_i}, P'_i) = \lambda(\vec{x_i}, e(\vec{x_i}, \vec{x'_i}, P'_i), P'_i)$. The assumption that $\lambda(\vec{x_i}, \vec{x'_i}, P_i) = \lambda(\vec{x_i}, \vec{x'_i}, P'_i)$ thus implies that

$$\lambda(\vec{x_i}, e(\vec{x_i}, \vec{x'_i}, P_i), P_i) = \lambda(\vec{x_i}, e(\vec{x_i}, \vec{x'_i}, P'_i), P'_i) \tag{26}$$

As $e(\vec{x_i}, \vec{x_i'}, P_i') \in \iota(\vec{x_i})$, it follows that the welfare gains from switching from $e(\vec{x_i}, \vec{x_i'}, P_i')$ to $\vec{x_i}$ the same for all $P \in D$, and hence

$$\lambda(\vec{x_i}, e(\vec{x_i}, \vec{x_i'}, P_i'), P_i') = \lambda(\vec{x_i}, e(\vec{x_i}, \vec{x_i'}, P_i'), P_i) \tag{27}$$

Combining with the previous equation yields

$$\lambda(\vec{x_i}, e(\vec{x_i}, \vec{x_i'}, P_i), P_i) = \lambda(\vec{x_i}, e(\vec{x_i}, \vec{x_i'}, P_i'), P_i) \tag{28}$$

As $\lambda(\vec{x_i}, \cdot, P_i)$ is a representation of the preferences $P_i$ and $e(\vec{x_i}, \vec{x_i'}, P_i), e(\vec{x_i}, \vec{x_i'}, P_i') \in \iota(\vec{x_i})$, the finding in Lemma 1 that $\iota(\vec{x_i})$ is ranked implies that

$$e(\vec{x_i}, \vec{x_i'}, P_i) = e(\vec{x_i}, \vec{x_i'}, P_i') \tag{29}$$

By equation 7,

$$\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}' \Leftrightarrow \sum_{i=1}^{I} u(\vec{x_i}|P_i) \leq \sum_{i=1}^{I} u(\vec{x_i'}|P_i) = \sum_{i=1}^{I} u(e(\vec{x_i}, \vec{x_i'}, P_i)|P_i) \Leftrightarrow$$

$$\sum_{i=1}^{I} (\mu_{\iota(\vec{x_i})}(\vec{x_i}) + \gamma(\iota(\vec{x_i})|P_i)) \leq \sum_{i=1}^{I} (\mu_{\iota(\vec{x_i})}(e(\vec{x_i}, \vec{x_i'}, P_i)) + \gamma(\iota(\vec{x_i})|P_i)) \Leftrightarrow$$

$$\sum_{i=1}^{I} \mu_{\iota(\vec{x_i})}(\vec{x_i}) \leq \sum_{i=1}^{I} \mu_{\iota(\vec{x_i})}(e(\vec{x_i}, \vec{x_i'}, P_i))$$

Similarly, $\mathbf{x} \preceq_{\mathbf{P'}} \mathbf{x}'$ if and only if $\sum_{i=1}^{I} \mu_{\iota(\vec{x_i})}(\vec{x_i}) \leq \sum_{i=1}^{I} \mu_{\iota(\vec{x_i})}(e(\vec{x_i}, \vec{x_i'}, P_i'))$. By equation 29, this is the same condition and hence $\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}'$ if and only if $\mathbf{x} \preceq_{\mathbf{P'}} \mathbf{x}'$, concluding the proof that the IIV axiom is satisfied. As the constitution represented by equation 7 satisfies the theorem's axioms, a constitution exists.

To prove uniqueness, define a function $z : X^2 \times D^2 \mapsto X$ as

$$z(\vec{x}, \vec{x'}, P, P') = e(\vec{x}, e(\vec{x'}, \vec{x}, P), P') \tag{30}$$

and, for $k < I$ and $(P_1, ..., P_k) \in D^k$, define the (partial) ranking $\preceq_{(P_1,...,P_k)}$ on $X^k$ as

$$(\vec{x_1}, ..., \vec{x_k}) \preceq_{(P_1,...,P_k)} (\vec{x_1'}, ..., \vec{x_k'}) \Leftrightarrow \forall \vec{x}_{k+1}, ..., \vec{x_I} \in X, P_{k+1}, ..., P_I \in D,$$

$$(\vec{x_1},...,\vec{x_k},\vec{x}_{k+1},...,\vec{x_I}) \preceq_{(P_1,...,P_k,P_{k+1},...,P_I)} (\vec{x'_1},...,\vec{x'_k},\vec{x}_{k+1},...,\vec{x_I})$$

**Claim 1.** *Let there be $P, P' \in D$ such that $\vec{x} \preceq_P \vec{x'}$ and $\vec{x'} \prec_{P'} \vec{x}$. Then, $z(\vec{x},\vec{x'},P,P') \prec_P \vec{x} \prec_P z(\vec{x},\vec{x'},P',P)$.*

*Proof.* Given that $\vec{x} \preceq_P \vec{x'}$ and $\vec{x'} \prec_{P'} \vec{x}$, it holds that $e(\vec{x'},\vec{x},P) \sim_P \vec{x} \preceq_P \vec{x'}$ and $\vec{x'} \prec_{P'} \vec{x} \sim_{P'} e(\vec{x'},\vec{x},P')$. As $e(\vec{x'},\vec{x},P'), e(\vec{x'},\vec{x},P) \in \iota(\vec{x'})$ and $\iota(\vec{x'})$ is Pareto-ranked, it follows that

$$e(\vec{x'},\vec{x},P) \prec_P e(\vec{x'},\vec{x},P') \text{ and } e(\vec{x'},\vec{x},P) \prec_{P'} e(\vec{x'},\vec{x},P') \tag{31}$$

Hence,

$$z(\vec{x},\vec{x'},P,P') = e(\vec{x},e(\vec{x'},\vec{x},P),P') \prec_{P'} e(\vec{x},e(\vec{x'},\vec{x},P'),P') = \vec{x} \tag{32}$$

Similarly,

$$\vec{x} = e(\vec{x},e(\vec{x'},\vec{x},P),P) \prec_P e(\vec{x},e(\vec{x'},\vec{x},P'),P) = z(\vec{x},\vec{x'},P',P)$$

The claim follows as $z(\vec{x},\vec{x'},P',P), \vec{x}, z(\vec{x},\vec{x'},P,P') \in \iota(\vec{x})$ and $\iota(\vec{x})$ is Pareto-ranked. □

**Claim 2.**

1. *For any $\vec{x},\vec{x'} \in X$, $P_1, P_2, P'_1, P'_2 \in D$ and $\vec{x_1},\vec{x}_2 \in \iota(\vec{x})$, $(\vec{x_1},\vec{x}_2) \sim_{(P'_1,P'_2)} (z(\vec{x_1},\vec{x'},P_1,P_2), z(\vec{x}_2,\vec{x'},P_2,P_1))$.*
2. *$\mu_{\iota(\vec{x})}(z(\vec{x},\vec{x'},P,P')) - \mu_{\iota(\vec{x})}(\vec{x}) = -(\mu_{\iota(\vec{x})}(z(\vec{x},\vec{x'},P',P)) - \mu_{\iota(\vec{x})}(\vec{x}))$.*
3. *For any $\vec{x''} \in \iota(\vec{x})$, it holds that $\mu_{\iota(\vec{x''})}(z(\vec{x''},\vec{x'},P,P')) - \mu_{\iota(\vec{x''})}(\vec{x''}) = \mu_{\iota(\vec{x})}(z(\vec{x},\vec{x'},P,P')) - \mu_{\iota(\vec{x})}(\vec{x})$.*

*Proof.* The proof of the first part of the claim is provided in the main text (the sketch of the proof of uniqueness). To prove the second and third parts, note that, as $e(\vec{x'},\vec{x},P) \sim_P \vec{x}$ and $e(\vec{x'},\vec{x},P) \in \iota(\vec{x'})$,

$$\mu_{\iota(\vec{x})}(\vec{x}) + \gamma(\iota(\vec{x})|P) = u(\vec{x}|P) = u(e(\vec{x'},\vec{x},P)|P) = \mu_{\iota(\vec{x'})}(e(\vec{x'},\vec{x},P)) + \gamma(\iota(\vec{x'})|P)$$

$$\Rightarrow \mu_{\iota(\vec{x'})}(e(\vec{x'},\vec{x},P)) = \mu_{\iota(\vec{x})}(\vec{x}) + \gamma(\iota(\vec{x})|P) - \gamma(\iota(\vec{x'})|P) \tag{33}$$

Similarly, as $u(e(\vec{x'}, \vec{x}, P)|P') = u(z(\vec{x}, \vec{x'}, P, P')|P')$ and $z(\vec{x}, \vec{x'}, P, P') \in \iota(\vec{x})$, it holds that

$$\mu_{\iota(\vec{x})}(z(\vec{x}, \vec{x'}, P, P')) + \gamma(\iota(\vec{x})|P') = \mu_{\iota(\vec{x})}(e(\vec{x'}, \vec{x}, P)) + \gamma(\iota(\vec{x'})|P') \tag{34}$$

Substituting in the previous equation yields

$$\mu_{\iota(\vec{x})}(z(\vec{x}, \vec{x'}, P, P')) + \gamma(\iota(\vec{x})|P') = (\mu_{\iota(\vec{x})}(\vec{x}) + \gamma(\iota(\vec{x})|P) - \gamma(\iota(\vec{x'})|P)) + \gamma(\iota(\vec{x'})|P')$$

Rearranging yields

$$\mu_{\iota(\vec{x})}(z(\vec{x}, \vec{x'}, P, P')) - \mu_{\iota(\vec{x})}(\vec{x}) = (\gamma(\iota(\vec{x})|P) - \gamma(\iota(\vec{x'})|P)) - (\gamma(\iota(\vec{x})|P') - \gamma(\iota(\vec{x'})|P'))$$

The second part of the claim follows by switching $P$ and $P'$. The third part of the claim follows because if $\vec{x''} \in \iota(\vec{x})$ then, by Lemma 1, $\iota(\vec{x}) = \iota(\vec{x''})$ and hence $\gamma(\iota(\vec{x})|P) = \gamma(\iota(\vec{x''})|P)$ and $\gamma(\iota(\vec{x})|P') = \gamma(\iota(\vec{x''})|P')$. □

**Claim 3.** *For each $\mathbf{P} \in D^I$, $\vec{x} \in \mathbb{R}^J_+$ and allocations $\mathbf{x}, \mathbf{x}' \in \iota(\vec{x})^I$, it holds that if $\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}'$ then $\mathbf{x} \preceq_{\mathbf{P}'} \mathbf{x}'$ for every $\mathbf{P}' \in D^I$.*

*Proof.* Let $\mathbf{x}, \mathbf{x}' \in \iota(\vec{x})^I$. By definition of $\iota(\vec{x})$, it holds that, for each $\mathbf{P}, \mathbf{P}' \in D^I$, $\lambda(\vec{x}_i, \vec{x'}_i, P_i) = \lambda(\vec{x}_i, \vec{x'}_i, P'_i)$. Hence, by the IIV axiom, $\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}'$ if and only if $\mathbf{x} \preceq_{\mathbf{P}'} \mathbf{x}'$. □

**Claim 4.** *For each $\mathbf{P} \in D^I$ and permutation $\pi : \{1, ..., I\} \mapsto \{1, ..., I\}$, it holds that if $(\vec{x_1}, ..., \vec{x_I}) \in \iota(\vec{x})$ then $(\vec{x_1}, ..., \vec{x_I}) \sim_{\mathbf{P}} (\vec{x}_{\pi(1)}, ..., \vec{x}_{\pi(I)})$.*

*Proof.* Weak Anonymity implies that, for each $P \in D$, $(\vec{x_1}, ..., \vec{x_I}) \sim_{(P,...,P)} (\vec{x}_{\pi(1)}, ..., \vec{x}_{\pi(I)})$ (as any two individuals with the same preferences must be treated symmetrically). By Claim 3, it follows that $(\vec{x_1}, ..., \vec{x_I}) \sim_{\mathbf{P}} (\vec{x}_{\pi(1)}, ..., \vec{x}_{\pi(I)})$ for every $\mathbf{P} \in D^I$. □

**Claim 5.** *For each $\vec{x} \in X$, $\mathbf{x} \in \iota(\vec{x})^I$ and $\mathbf{P} \in D^I$, it holds that*

$$\mathbf{x} \sim_{\mathbf{P}} (\mu^{-1}_{\iota(\vec{x})}(\frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i})), ..., \mu^{-1}_{\iota(\vec{x})}(\frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i})))$$

*Proof.* Define $N : \iota(\vec{x})^I \mapsto \{0, ..., I\}$ so that $N(\mathbf{x})$ is the number of elements in $\mu_{\iota(\vec{x})}(\vec{x_1}), ..., \mu_{\iota(\vec{x})}(\vec{x_I})$ which are different from $\frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i})$. The claim trivially holds for all $\mathbf{x} \in \iota(\vec{x})^I$ for which $N(\mathbf{x}) = 0$. The value $N(\mathbf{x}) = 1$ is not possible because there cannot be only one element that is different from the average – thus, the claim trivially holds for all $\mathbf{x}$ for which $N(\mathbf{x}) = 1$ (which is an empty set).

Assume that the claim holds for any $\mathbf{x} \in \iota(\vec{x})^I$ for which $N(\mathbf{x}) \geq k$, where $1 \leq k < I$. Let there be $\mathbf{x} \in X^I$ for which $N(\mathbf{x}) = k + 1$. By Claim 4, I can assume without loss of generality that $\mu_{\iota(\vec{x})}(\vec{x_1}) < \frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i}) < \mu_{\iota(\vec{x})}(\vec{x}_2)$.

As $|D| > 1$, let there be $P, P' \in D$ such that $P \neq P'$. This implies that there exists $\vec{x'} \in X$ such that $\vec{x} \preceq_P \vec{x'}$ and $\vec{x'} \prec_{P'} \vec{x}$ (this follows because $\lambda(\vec{x}, \cdot, P)$ represents $P$ and $\lambda(\vec{x}, \cdot, P')$ represents $P'$). By Claim 1,

$$z(\vec{x}, \vec{x'}, P, P') \prec_{P'} \vec{x} \tag{35}$$

As $\mu_{\iota(\cdot)}(\cdot)+\gamma(\iota(\cdot)|P')$ represents the preferences $P'$, it follows that $\mu_{\iota(\vec{x})}(z(\vec{x}, \vec{x'}, P, P')) < \mu_{\iota(\vec{x})}(\vec{x})$. By the second part of Claim 2, it follows that $\mu_{\iota(\vec{x})}(z(\vec{x}, \vec{x'}, P', P)) - \mu_{\iota(\vec{x})}(\vec{x}) > 0$.

Note that $z(\vec{x}, \vec{x}, P', P) = \vec{x}$, and that $z(\vec{x}, (1-\eta)\vec{x} + \eta\vec{x'}, P', P)$ is a continuous function of $\eta$ (this follows from the assumption that preferences in $D$ are continuous). As $\mu_{\iota(\vec{x})}(\cdot)$ is continuous, it follows that there exists $\eta \in [0, 1]$ such that, for some integer $m$,

$$\mu_{\iota(\vec{x})}(z(\vec{x}, (1-\eta)\vec{x} + \eta\vec{x'}, P', P)) - \mu_{\iota(\vec{x})}(\vec{x}) = \frac{\frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i}) - \mu_{\iota(\vec{x})}(\vec{x_1})}{m} \tag{36}$$

Define sequences $\{\vec{r}_n\}_{n=0}^{m}, \{\vec{s}_n\}_{n=0}^{m} \subset \iota(\vec{x})$ as follows: $\vec{r}_0 = \vec{x_1}$, $\vec{s}_0 = \vec{x}_2$, and, for $n \geq 0$, $\vec{r}_{n+1} = z(\vec{r}_n, (1-\eta)\vec{x} + \eta\vec{x'}, P', P)$ and $\vec{s}_{n+1} = z(\vec{s}_n, (1-\eta)\vec{x} + \eta\vec{x'}, P, P')$.

Note that, for every $n$, $\vec{r}_n, \vec{s}_n \in \iota(\vec{x})$ (this is easily established by induction given that this holds for $n = 0$). By the first part of Claim 2, for every $\mathbf{P} \in D^I$, it holds that $(\vec{r}_{n+1}, \vec{s}_{n+1}, \vec{x}_3, ..., \vec{x_I}) \sim_{\mathbf{P}} (\vec{r}_n, \vec{s}_n, \vec{x}_3, ..., \vec{x_I})$. Hence, by induction, it holds that

$$(\vec{r}_m, \vec{s}_m, \vec{x}_3, ..., \vec{x_I}) \sim_{\mathbf{P}} \mathbf{x} \tag{37}$$

for every $\mathbf{P} \in D^I$.

Further, note that, by the third clause of Claim 2 (as $\iota(\vec{x}) = \iota(\vec{r}_n)$ for all $n$),

$$\mu_{\iota(\vec{x})}(\vec{r}_{n+1}) - \mu_{\iota(\vec{x})}(\vec{r}_n) = \mu_{\iota(\vec{x})}(z(\vec{r}_n, (1-\eta)\vec{x} + \eta\vec{x'}, P', P)) - \mu_{\iota(\vec{x})}(\vec{r}_n) = \tag{38}$$

$$\mu_{\iota(\vec{x})}(z(\vec{x}, (1-\eta)\vec{x} + \eta\vec{x'}, P', P)) - \mu_{\iota(\vec{x})}(\vec{x}) = \frac{\frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i}) - \mu_{\iota(\vec{x})}(\vec{x_1})}{m}$$

(where the last equality follows from equation 36).

Consequently,

$$\mu_{\iota(\vec{x})}(\vec{r}_m) - \mu_{\iota(\vec{x})}(\vec{r}_0) = \frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i}) - \mu_{\iota(\vec{x})}(\vec{x_1})$$

As $\vec{r}_0 = \vec{x_1}$, it follows that

$$\mu_{\iota(\vec{x})}(\vec{r}_m) = \frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i})$$

Thus, $N(\vec{r}_m, \vec{s}_m, \vec{x}_3, ..., \vec{x_I}) \leq N(\mathbf{x}) - 1 \leq k$. By the induction hypothesis, for every $\mathbf{P} \in D^I$, social preferences are indifferent between the allocation $(\vec{r}_m, \vec{s}_m, \vec{x}_3, ..., \vec{x_I})$ and an allocation in which all individuals get the bundle $\vec{y} = \mu^{-1}_{\iota(\vec{x})}((\mu_{\iota_{\vec{x}}}(\vec{r}_m) + \mu_{\iota_{\vec{x}}}(\vec{s}_m) + \sum_{i=3}^{I} \mu_{\iota_{\vec{x}}}(\vec{x_i}))/I)$. Note that, by equation 38,

$$\mu_{\iota(\vec{x})}(\vec{r}_{n+1}) = \mu_{\iota(\vec{x})}(\vec{r}_n) + \frac{\frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i}) - \mu_{\iota(\vec{x})}(\vec{x_1})}{m} \tag{39}$$

and, similarly (using the second part of Claim 2)

$$\mu_{\iota(\vec{x})}(\vec{s}_{n+1}) = \mu_{\iota(\vec{x})}(\vec{s}_n) - \frac{\frac{1}{I}\sum_{i=1}^{I} \mu_{\iota(\vec{x})}(\vec{x_i}) - \mu_{\iota(\vec{x})}(\vec{x_1})}{m} \tag{40}$$

It thus follows inductively that $\mu_{\iota(\vec{x})}(\vec{r}_m) + \mu_{\iota(\vec{x})}(\vec{s}_m) = \mu_{\iota(\vec{x})}(\vec{r}_0) + \mu_{\iota(\vec{x})}(\vec{s}_0) = \mu_{\iota(\vec{x})}(\vec{x_1}) + \mu_{\iota(\vec{x})}(\vec{x}_2)$, and hence $\vec{y} = \mu^{-1}_{\iota(\vec{x})}(\sum_{i=1}^{I} \mu_{\iota_{\vec{x}}}(\vec{x_i})/I)$. By expression 37, it thus follows that the claim holds for $\mathbf{x}$. This concludes the proof that the claim holds for all $\mathbf{x} \in \iota(\vec{x})^I$ such that $N(\mathbf{x}) = k + 1$.

By induction, it follows that the claim holds for all $\mathbf{x} \in \iota(\vec{x})^I$ for which $N(\mathbf{x}) \leq I$ – which is the entire set $\iota(\vec{x})^I$. □

To conclude the proof of uniqueness, note that, by the Pareto principle, it must hold that, for every $\mathbf{x} \in X^I$, $\vec{x} \in X$ and $\mathbf{P} \in D^I$, $\mathbf{x} \sim_{\mathbf{P}} (e(\vec{x}, \vec{x_1}, P_1), ..., e(\vec{x}, \vec{x_I}, P_I))$. Thus, for each $\mathbf{x}' \in X^I$, it holds that

$$\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}' \Leftrightarrow (e(\vec{x}, \vec{x_1}, P_1), ..., e(\vec{x}, \vec{x_I}, P_I)) \preceq_{\mathbf{P}} (e(\vec{x}, \vec{x'_1}, P_1), ..., e(\vec{x}, \vec{x'_I}, P_I))$$

By the above claim,

$$(e(\vec{x},\vec{x_1},P_1),...,e(\vec{x},\vec{x_I},P_I)) \sim_{\mathbf{P}} (\mu^{-1}_{\iota(\vec{x})}(\frac{1}{I}\sum_{i=1}^{I}\mu_{\iota(\vec{x})}(e(\vec{x},\vec{x_i},P_i)),...,\mu^{-1}_{\iota(\vec{x})}(\frac{1}{I}\sum_{i=1}^{I}\mu_{\iota(\vec{x})}(e(\vec{x},\vec{x_i},P_i)))$$

Thus, as $\iota(\vec{x})$ is Pareto-ranked (and, given the restriction to increasing preferences, $\mu_{\iota(\vec{x})}$ must be strictly monotone), the Pareto principle requires that

$$\mathbf{x} \preceq_{\mathbf{P}} \mathbf{x}' \Leftrightarrow \frac{1}{I}\sum_{i=1}^{I}\mu_{\iota(\vec{x})}(e(\vec{x},\vec{x_i},P_i)) \leq \frac{1}{I}\sum_{i=1}^{I}\mu_{\iota(\vec{x})}(e(\vec{x},\vec{x'_i},P_i))$$

which concludes the proof of uniqueness.

**Proof of the second part of the theorem.** It is useful to introduce the following claim:

**Claim 6.** *For every $P,P' \in D$ and $\vec{x},\vec{x'} \in X$, it holds that if $\vec{y} = z(\vec{x},\vec{x'},P,P')$ then $\vec{x} = z(\vec{y},\vec{x'},P',P)$.*

*Proof.* Let $\vec{x''} = e(\vec{x'},\vec{x},P)$. By definition, $\vec{y} = z(\vec{x},\vec{x'},P,P') = e(\vec{x},\vec{x''},P')$, and hence $\vec{y} \sim_{P'} \vec{x''}$

Note that $z(\vec{y},\vec{x'},P',P) = e(\vec{y},e(\vec{x'},\vec{y},P'),P)$. As $\vec{x''} \in \iota(\vec{x'})$ and $\vec{y} \sim_{P'} \vec{x''}$, it follows that $e(\vec{x'},\vec{y},P') = \vec{x''}$ and hence $z(\vec{y},\vec{x'},P',P) = e(\vec{y},\vec{x''},P)$. By definition of $\vec{x''}$, it holds that $\vec{x''} \sim_P \vec{x}$, and, as $\vec{x} \in \iota(\vec{y})$, it follows that $\vec{x} = e(\vec{y},\vec{x''},P)$ and hence $\vec{x} = z(\vec{y},\vec{x'},P',P)$.

□

Assume that the $\{\preceq_{\mathbf{P}}\}_{\mathbf{P}\in D^I}$ is a constitution that satisfies the theorem's axioms. Fix some $P,P' \in D$ and $\vec{x},\vec{x'} \in X$ such that $\vec{x} \in \iota(\vec{1})$, $\vec{x} \preceq_P \vec{x'}$ and $\vec{x'} \prec_{P'} \vec{x}$. Define a sequence of functions $\{\phi_n : \iota(\vec{x}) \mapsto \mathbb{R}\}_{n=0}^{\infty}$ as follows. To define $\phi_0$, I define a sequence $\{\vec{x}_m\}_{m=-\infty}^{\infty} \subset \iota(\vec{x})$ as follows: $\vec{x_0} = \vec{1} = e(\vec{x},\vec{1},P)$; for $m > 0$, $\vec{x}_m = z(\vec{x}_{m-1},\vec{x'},P',P)$, and, for $m < 0$, $\vec{x}_m = z(\vec{x}_{m+1},\vec{x'},P,P')$.

Note that, by Claim 6, for every $m$ it holds that

$$\vec{x_m} = z(\vec{x}_{m+1},\vec{x'},P,P') = z(\vec{x}_{m-1},\vec{x'},P',P) \tag{41}$$

**Claim 7.** *For very $m$, $\vec{x}_m \prec_P \vec{x}_{m+1}$.*

*Proof.* For $m = 0$, this follows from Claim 1. Assume that this holds for $m \geq 0$. By the first clause of Claim 2 (using equation 41),

$$(\vec{x}_m, \vec{x}_{m+2}) \sim_{(P,P)} (z(\vec{x}_m, \vec{x'}, P', P), z(\vec{x}_{m+2}, \vec{x'}, P, P') = (\vec{x}_{m+1}, \vec{x}_{m+1})$$

By the induction hypothesis, $\vec{x}_m \prec_P \vec{x}_{m+1}$ and hence the Pareto condition requires that $\vec{x}_{m+1} \prec_P \vec{x}_{m+2}$ – concluding the proof that this holds for every $m \geq 0$. Similar steps can be used to establish that this holds for $m < 0$. □

**Claim 8.** *The sequence* $\{\vec{x}_m\}_{m=-\infty}^{\infty}$ *is unbounded (in* $\mathbb{R}_+^J$*).*

*Proof.* Assume by way of contradiction that there is an upper-bound on $\{\vec{x}_m\}_{m=-\infty}^{\infty}$ (the proof for the case of a lower-bound is similar and hence omitted). As $\{\vec{x}_m\}_{m=-\infty}^{\infty}$ is an increasing and bounded sequence, it converges - let $\vec{x^*}$ be its limit. By continuity of $\vec{z}$, it follows that

$$\vec{x^*} = z(\vec{x^*}, \vec{x'}, P, P')$$

By the first clause of Claim 2,

$$(\vec{x}_0, \vec{x^*}) \sim_{(P,P)} (z(\vec{x}_0, \vec{x'}, P', P), z(\vec{x^*}, \vec{x'}, P, P')) = (\vec{x}_1, \vec{x^*})$$

Given Claim 7, this is a contradiction to the Pareto principle. It thus follows that $\{\vec{x}_m\}_{m=-\infty}^{\infty}$ is unbounded. □

Define $\phi_0(\vec{x}_m) = m$, and, for $\vec{r} \in \iota(\vec{x})$ such that $\vec{x}_m \prec_P \vec{r} \prec_P \vec{x}_{m+1}$, let $\phi_0(\vec{r})$ take a value between $m$ and $m+1$ (note that, by Claim 8, there exist such $m$ for every $\vec{r} \in \iota(\vec{x})$). it is straightforward to establish that it is possible to specify $\phi_0$ to be strictly increasing and continuous.

Define $\{\vec{x}_{0,m}\}_{m=-\infty}^{\infty} = \{\vec{x}_m\}_{m=-\infty}^{\infty}$. Assume that $n$ is such that $\vec{x}_{n,0} = \vec{x}_{0,0}$, and

1. For every $m$, $\vec{x}_{n,m} \prec_P \vec{x}_{n,m+1}$.

2. There exists $\vec{y}_n \in X$ such that, for every $m$, $\vec{x}_{n,m} = z(\vec{x}_{n,m+1}, \vec{y}_n, P, P')$.

3. $\phi_n(\vec{x}_{n,m}) = m/2^n$.

4. $\phi_n$ is strictly increasing and continuous.

5. for every $n' < n$, $\phi_n(\vec{x}_{n',m}) = \phi_{n'}(\vec{x}_{n',m})$.

I have established that these properties hold for $n = 0$ (the last property, trivially since $n \geq 0$).

Define the sequence $\{\vec{x}_{n+1,m}\}_{m=-\infty}^{\infty}$ as follows. Let $\vec{x}_{n+1,0} = \vec{x}_{0,0}$. Define the vectors $\vec{x}_{n+1,1}$ and $\vec{y}_{n+1}$ as a solution to the equations

$$\vec{x}_{n+1,1} = z(\vec{x}_{n,0}, \vec{y}_{n+1}, P', P) \text{ and } \vec{x}_{n,1} = z(\vec{x}_{n+1,1}, \vec{y}_{n+1}, P', P) \tag{42}$$

To establish that a solution exists, combine the two equations into an equation in which $\vec{y}_{n+1}$ is the single unknown:

$$\vec{x}_{n,1} = z(z(\vec{x}_{n,0}, \vec{y}_{n+1}, P', P), \vec{y}_{n+1}, P', P) \tag{43}$$

Note that, for $\vec{y}_{n+1} = \vec{x}_{n,0}$, the right hand side takes the value $\vec{x}_{n,0}$ – as $\vec{x}_{n,0} \prec_P \vec{x}_{n,1}$, this cannot be a solution. Alternatively, for $\vec{y}_{n+1} = \vec{y}_n$, it holds that $z(\vec{x}_{n,0}, \vec{y}_{n+1}, P', P) = \vec{x}_{n,1}$ and hence the right hand side takes the value $\vec{x}_{n,2}$: for this choice of $\vec{y}_{n+1}$, the left hand side is strictly preferred over the right hand side (by the first property listed above). By the continuity of individual preferences, there exists $\eta \in (0, 1)$ such that $\vec{y}_{n+1} = (1 - \eta)\vec{x}_{n,0} + \eta\vec{y}_n$ is a solution to the above equation.

Define $\{\vec{x}_{n+1,m}\}_{m=-\infty}^{\infty}$ to be consistent with the second property, given this choice of $\vec{y}_{n+1}$. Following similar steps to Claim 7, the first property follows (note that the application of Claim 1 for $m = 0$ at the beginning of the proof of Claim 7 requires that $\vec{x}_{0,0} \preceq_P \vec{y}_{n+1}$ and $\vec{y}_{n+1} \prec_{P'} \vec{y}_{n+1}$; this follows because, otherwise, $z(\vec{x}_{n+1,1}, \vec{y}_{n+1}, P', P) \prec_P \vec{x}_{n,0}$, which is a contradiction to the assumption that $\vec{x}_{n,1} = z(\vec{x}_{n+1,1}, \vec{y}_{n+1}, P', P)$ and $\vec{x}_{n,0} \prec_P \vec{x}_{n,1}$).

Define $\phi_{n+1}(\vec{x}_{n+1,m}) = m/2^{n+1}$, to be consistent with the third property, and extrapolate $\phi_n$ to be an increasing and continuous function from $\iota(\vec{x})$ to $\mathbb{R}$ (this guarantees consistency with the fourth property).

To establish that $\phi_n$ is consistent with the fifth property, I establish that $\vec{x}_{n+1,2m} = \vec{x}_{n,m}$ (the fifth property trivially follows by backward induction, as $\phi_{n+1}(\vec{x}_{n+1,2m}) = 2m/2^{n+1} = m/2^n = \phi_n(\vec{x}_{n,m})$). To see this, note that, by construction, this property holds for $m = 0$ and $m = 1$. I use induction to show that this holds for every $m$. Assume that this holds for $m - 1$ and $m$. To establish that is holds for $m + 1$, note that, by the first clause of Claim 2,

$$(\vec{x}_{n,m-1}, \vec{x}_{n,m+1}) \sim_{(P,P)} (\vec{x}_{n,m}, \vec{x}_{n,m})$$

and, similarly,

$$(\vec{x}_{n+1,2(m-1)}, \vec{x}_{n+1,2(m+1)}) \sim_{(P,P)} (\vec{x}_{n+1,2m-1}, \vec{x}_{n+1,2m+1}) \sim_{(P,P)} (\vec{x}_{n+1,2m}, \vec{x}_{n+1,2m})$$

as, by the induction hypothesis, $(\vec{x}_{n+1,2m}, \vec{x}_{n+1,2m}) = (\vec{x}_{n,m}, \vec{x}_{n,m})$, it follows that

$$(\vec{x}_{n,m-1}, \vec{x}_{n,m+1}) \sim_{(P,P)} (\vec{x}_{n+1,2(m-1)}, \vec{x}_{n+1,2(m+1)})$$

As, by the induction hypothesis $\vec{x}_{n,m-1} = \vec{x}_{n+1,2(m-1)}$, the Pareto principle implies that $\vec{x}_{n,m+1} \sim_P \vec{x}_{n+1,2(m+1)}$. As $\vec{x}_{n,m+1}, \vec{x}_{n+1,2(m+1)} \in \iota(\vec{x})$, this implies that $\vec{x}_{n,m+1} = \vec{x}_{n+1,2(m+1)}$.

A similar proof establishes that, if $\vec{x}_{n+1,2(m+1)} = \vec{x}_{n,m+1}$ and $\vec{x}_{n+1,2m} = \vec{x}_{n,m}$ then $\vec{x}_{n+1,2(m-1)} = \vec{x}_{n,m-1}$ (which is necessary to show for $m \leq 0$). For the sake of brevity the proof is omitted.

Note that, for each $\vec{r} \in \iota(\vec{1})$, the sequence $\{\phi_n(\vec{r})\}_{n=0}^{\infty}$ is a Cauchy sequence: for each $\epsilon > 0$, there exists $N$ such that for each $n, k > N$, it holds that $|\phi_n(\vec{r}) - \phi^k(\vec{r})| < \epsilon$. To see this, choose $N$ such that $1/2^N < \epsilon$, and let $m$ be such that $x_{N,m} \prec_P \vec{r} \prec_P \vec{x}_{N,m+1}$ (note that, by Claim 8 and the fifth property, it is straightforward to establish that such an $m$ exists for every $N$). Given the definition of $\phi_N$ as a strictly increasing function, it holds that $\phi_N(\vec{r}) \in (\phi_N(\vec{x}_{N,m}), \phi_N(\vec{x}_{N,m+1})) = (m/2^N, (m+1)/2^N)$. Given the fifth property, it holds that, for every $n > N$, $(\phi_N(\vec{x}_{N,m}), \phi_N(\vec{x}_{N,m+1})) = (\phi_n(\vec{x}_{N,m}), \phi_n(\vec{x}_{N,m+1}))$ and hence $\phi_n(\vec{r}) \in (m/2^N, (m+1)/2^N)$. It follows that, if $k, n > N$, then $\phi_n(\vec{r}), \phi_k(\vec{r}) \in (m/2^N, (m+1)/2^N)$ and hence $|\phi_n(\vec{r}) - \phi_k(\vec{r})| < 1/2^N < \epsilon$.

As, for each $\vec{r} \in \iota(\vec{1})$, $\{\phi_n(\vec{r})\}_{n=0}^{\infty}$ is a Cauchy sequence, it follows that $\lim_{n\to\infty} \phi_n$ exists. Define $\mu_{\iota(\vec{1})} = \lim_{n\to\infty} \phi_n$.

To establish that $\mu_{\iota(\vec{1})}$ is continuous, fix some $\vec{r} \in \iota(\vec{1})$ and $\epsilon > 0$. To establish that there exists $\delta > 0$ such that, if $||\vec{r} - \vec{r'}|| < \delta$ then $|\mu_{\iota(\vec{1})}(\vec{r}) - \mu_{\iota(\vec{1})}(\vec{r'})| < \epsilon$, let $N$ be such that $1/2^N < \epsilon$ and observe that if, for some $m$, $\vec{x}_{N,m} \prec_P \vec{r}, \vec{r'} \prec_P \vec{x}_{N,m+1}$, then, for each $n > N$, it holds that $|\phi_n(\vec{r}) - \phi_n(\vec{r'})| < 1/2^N < \epsilon$. Hence, in the limit, $|\mu_{\iota(\vec{1})}(\vec{r}) - \mu_{\iota(\vec{1})}(\vec{r'})| < \epsilon$, concluding the proof of continuity.

To establish that the limit function $\mu_{\iota(\vec{1})}$ is strictly monotone, note that, if $\vec{r} \prec_P \vec{r'}$, then there exists $N$ and $m$ such that $\vec{r} \prec_P \vec{x}_{N,m+1}$ and $\vec{x}_{N,m+1} \prec_P \vec{r'}$. It is straightforward to see that, for every $n > N$, $\phi_n(\vec{r'}) - \phi_n(\vec{r}) > 1/2^N$ and hence, in the limit, $\mu_{\iota(\vec{1})}(\vec{r'}) - \mu_{\iota(\vec{1})}(\vec{r}) \geq 1/2^N > 0$. It follows that $\mu_{\iota(\vec{1})}$ is strictly monotone.

**Claim 9.** *For each* $\vec{r}, \vec{r'}, \vec{r''} \in \iota(\vec{1})$ *such that* $(\vec{1}, \vec{r}) \sim_{(P,P')} (\vec{r'}, \vec{r''})$*, it holds that* $\mu_{\iota(\vec{1})}(\vec{r}) = \mu_{\iota(\vec{1})}(\vec{r'}) + \mu_{\iota(\vec{1})}(\vec{r''})$.

*Proof.* Let $\vec{r}, \vec{r'}, \vec{r''} \in \iota(\vec{1})$ be such that $(\vec{1}, \vec{r}) \sim_{(P,P')} (\vec{r'}, \vec{r''})$.

Let $\{m(n)\}_{n=0}^{\infty} \subset \mathbb{Z}$ be a sequence of integers such that $\lim_{n\to\infty} m(n)/2^n = \mu_{\iota(\vec{1})}(\vec{r})$, and let $\{m'(n)\}_{n=0}^{\infty} \subset \mathbb{Z}$ be a sequence of integers such that $\lim_{n\to\infty} m'(n)/2^n = \mu_{\iota(\vec{1})}(\vec{r'})$.

I establish that, for every $n$,

$$(\vec{1}, \mu_{\iota(\vec{1})}^{-1}(\frac{m(n)}{2^n})) \sim_{(P,P')} (\mu_{\iota(\vec{1})}^{-1}(\frac{m'(n)}{2^n}), \mu_{\iota(\vec{1})}^{-1}(\frac{m(n)-m'(n)}{2^n})) \tag{44}$$

I show this by induction on $k = m'(n)$. Without loss of generality, assume that $m'(n) \geq 0$ (the proof for the case $m'(n) \leq 0$ is analogous). For $k = 0$, this follows trivially from the reflexivity of the indifference relation. Assume that, for some $k \geq 0$,

$$(\vec{1}, \mu_{\iota(\vec{1})}^{-1}(\frac{m(n)}{2^n})) \sim_{(P,P')} (\mu_{\iota(\vec{1})}^{-1}(\frac{k}{2^n}), \mu_{\iota(\vec{1})}^{-1}(\frac{m(n)-k}{2^n})) \tag{45}$$

Note that, by the fifth property, $\mu_{\iota(\vec{1})}(\vec{x}_{n,m}) = \phi_n(\vec{x}_{n,m}) = m/2^n$; thus, the above condition can be rewritten as

$$(\vec{1}, \vec{x}_{n,m(n)}) \sim_{(P,P')} (\vec{x}_{n,k}, \vec{x}_{n,m(n)-k}) \tag{46}$$

Using the first part of Claim 2, it follows that

$$(\vec{x}_{n,k}, \vec{x}_{n,m(n)-k}) \sim_{P,P'} (z(\vec{x}_{n,k}, \vec{y}_n, P', P), z(\vec{x}_{n,m(n)-k}, \vec{y}_n, P, P')) \tag{47}$$

By the second property, $\vec{x}_{n,m(n)-k-1} = z(\vec{x}_{n,m(n)-k}, \vec{y}_n, P, P')$. Similarly, the second property implies that $\vec{x}_{n,k} = z(\vec{x}_{n,k+1}, \vec{y}_n, P, P')$, and, hence, by Claim 6, $\vec{x}_{n,k+1} = z(\vec{x}_{n,k}, \vec{y}_n, P', P)$. Substituting, the above expression can be rewritten as

$$(\vec{1}, \mu_{\iota(\vec{1})}^{-1}(\frac{m(n)}{2^n})) = (\vec{1}, \vec{x}_{n,m(n)}) \sim_{(P,P')} (\vec{x}_{n,k}, \vec{x}_{n,m(n)-k}) \sim_{P,P'} (\vec{x}_{n,k+1}, \vec{x}_{n,m(n)-k-1}) =$$

$$(\mu_{\iota(\vec{1})}^{-1}(\frac{k+1}{2^n}), \mu_{\iota(\vec{1})}^{-1}(\frac{m(n)-(k+1)}{2^n}))$$

Concluding the proof by induction and thus establishing that, for each $n$, expression 44 holds.

As $m(n)/2^n \rightarrow_{n\to\infty} \mu_{\iota(\vec{1})}(\vec{r})$ and $m'(n)/2^n \rightarrow_{n\to\infty} \mu_{\iota(\vec{1})}(\vec{r'})$, it follows that $(m(n) - m'(n))/2^n \rightarrow_{n\to\infty} \mu_{\iota(\vec{1})}(\vec{r}) - \mu_{\iota(\vec{1})}(\vec{r'})$. Denote $\vec{s} = \mu^{-1}_{\iota(\vec{1})}(\mu_{\iota(\vec{1})}(\vec{r}) - \mu_{\iota(\vec{1})}(\vec{r'}))$

Assume by way of contradiction that $\vec{s} \neq \vec{r''}$; in particular, without loss of generality, assume that $\mu_{\iota(\vec{1})}(\vec{s}) = \mu_{\iota(\vec{1})}(\vec{r}) - \mu_{\iota(\vec{1})}(\vec{r'}) < \mu_{\iota(\vec{1})}(\vec{r''})$ (the proof for the opposite inequality is similar).

Note that the only requirement for specifying the sequences $\{m(n)\}_{n=0}^{\infty}$ and $\{m'(n)\}_{n=0}^{\infty}$ was that $m(n)/2^n \rightarrow_{n\to\infty} \mu_{\iota(\vec{1})}(\vec{r})$ and $m'(n)/2^n \rightarrow_{n\to\infty} \mu_{\iota(\vec{1})}(\vec{r'})$. It is therefore possible to choose $\{m(n)\}_{n=0}^{\infty}$ such that $m(n)/2^n > \mu_{\iota(1)}(\vec{r})$ and $m'(n)/2^n < \mu_{\iota(1)}(\vec{r'})$ for every $n$.

As $\mu_{\iota(\vec{1})}$ is strictly monotone, the choice $m(n)/2^n > \mu_{\iota(1)}(\vec{r})$ implies that $\vec{r} \prec_{P'} \mu^{-1}_{\iota(\vec{1})}(\frac{m(n)}{2^n})$. Hence, by the Pareto principle,

$$(\vec{1}, \vec{r}) \prec_{(P,P')} (\vec{1}, \mu^{-1}_{\iota(\vec{1})}(\frac{m(n)}{2^n})) \sim_{(P,P')} (\mu^{-1}_{\iota(\vec{1})}(\frac{m'(n)}{2^n}), \mu^{-1}_{\iota(\vec{1})}(\frac{m(n)-m'(n)}{2^n}))$$

where the last indifference follows from expression 44. Similarly, as $\mu_{\iota(\vec{1})}$ is strictly monotone, the choice $m'(n)/2^n < \mu_{\iota(1)}(\vec{r'})$ implies that $\mu^{-1}_{\iota(\vec{1})}(\frac{m'(n)}{2^n}) \prec_{P'} \vec{r'}$ and, hence, by the Pareto principle,

$$(\mu^{-1}_{\iota(\vec{1})}(\frac{m'(n)}{2^n}), \mu^{-1}_{\iota(\vec{1})}(\frac{m(n)-m'(n)}{2^n})) \prec_{(P,P')} (\vec{r'}, \mu^{-1}_{\iota(\vec{1})}(\frac{m(n)-m'(n)}{2^n}))$$

As $(m(n)-m'(n))/2^n \rightarrow_{n\to\infty} \mu_{\iota(\vec{1})}(\vec{s}) < \mu_{\iota(\vec{1})}(\vec{r''})$, for $n$ sufficiently large, it holds that $(m(n) - m'(n))/2^n < \mu_{\iota(\vec{1})}(\vec{r''})$. As $\mu_{\iota(\vec{1})}$ is strictly increasing and continuous, so is its inverse, and hence $\mu^{-1}_{\iota(\vec{1})}(\frac{m(n)-m'(n)}{2^n}) \prec_{P'} \vec{r''}$; thus, by the Pareto principle, for $n$ sufficiently large it holds that

$$(\vec{r'}, \mu^{-1}_{\iota(\vec{1})}(\frac{m(n)-m'(n)}{2^n})) \prec_{(P,P')} (\vec{r'}, \vec{r''})$$

Combining, these indifference relations imply that

$$(\vec{1}, \vec{r}) \prec_{(P,P')} (\vec{r'}, \vec{r''})$$

in contradiction to the assumption that $(\vec{1}, \vec{r}) \sim_{(P,P')} (\vec{r'}, \vec{r''})$. This contradiction establishes that $\vec{s} = \vec{r''}$, and hence $\mu_{\iota(\vec{1})}(\vec{r}) - \mu_{\iota(\vec{1})}(\vec{r'}) = \mu_{\iota(\vec{1})}(\vec{s}) = \mu_{\iota(\vec{1})}(\vec{r''})$; rearranging, this implies that $\mu_{\iota(\vec{1})}(\vec{r}) = \mu_{\iota(\vec{1})}(\vec{r'}) + \mu_{\iota(\vec{1})}(\vec{r''})$, concluding the proof of the claim.

□

For each $\vec{r} \in X$, define the functions $\mu_{\iota(\vec{r})}$ and $\{\gamma(\iota(\vec{r})|P'')\}_{P'' \in D}$ as follows.

$$\mu_{\iota(\vec{r})}(\vec{r}) = \mu_{\iota(\vec{1})}(e(\vec{1}, \vec{r}, P)) \tag{48}$$

and, for each $P'' \in D$, define

$$\gamma(\iota(\vec{r})|P'') = \mu_{\iota(\vec{1})}(z(\vec{1}, \vec{r}, P, P'')) \tag{49}$$

(it is straightforward to establish that $\gamma(\iota(\vec{r})|P'')$ is well-defined, because if $\iota(\vec{r}) = \iota(\vec{r'})$ for some $\vec{r}, \vec{r'} \in X$, then $z(\vec{1}, \vec{r}, P, P'') = z(\vec{1}, \vec{r'}, P, P'')$).

**Claim 10.** *For each $\vec{r} \in X$ and $P'' \in D$, it holds that $\mu_{\iota(\vec{r})}(\vec{r}) + \gamma(\iota(\vec{r})|P'') = \mu_{\iota(\vec{1})}(e(\vec{1}, \vec{r}, P'')$.*

*Proof.* Note that

$$z(e(\vec{1}, \vec{r}, P''), \vec{r}, P'', P) = e(\vec{1}, e(e(\vec{1}, \vec{r}, P''), \vec{r}, P''), P) = e(\vec{1}, \vec{r}, P)$$

Thus, by the first clause of Claim 2, it holds that

$$(\vec{1}, e(\vec{1}, \vec{r}, P'')) \sim_{(P,P'')} (z(\vec{1}, \vec{r}, P, P''), z(e(\vec{1}, \vec{r}, P''), \vec{r}, P'', P)) = \tag{50}$$

$$(z(\vec{1}, \vec{r}, P, P''), e(\vec{1}, \vec{r}, P))$$

By Claim 3, it follows that

$$(\vec{1}, e(\vec{1}, \vec{r}, P'')) \sim_{(P,P')} (z(\vec{1}, \vec{r}, P, P''), e(\vec{1}, \vec{r}, P)) \tag{51}$$

By Claim 9, this implies that

$$\mu_{\iota(\vec{1})}(e(\vec{1}, \vec{r}, P'')) = \mu_{\iota(\vec{1})}(z(\vec{1}, \vec{r}, P, P'')) + \mu_{\iota(\vec{1})}(e(\vec{1}, \vec{r}, P)) \tag{52}$$

By the definitions of $\mu_{\iota(\vec{r})}$ and $\gamma(\iota(\vec{r})|P'')$ (equations 48 and 49), the claim follows. □

To conclude the proof of the theorem, note that $\mu_{\iota(\vec{1})}(e(\vec{1}, \cdot, P''))$ is a representation of the preferences $P''$. Thus, the above claim establishes that $\mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P'')$ is a representation of the preferences $P''$.

# F Proof of Corollary 1

Let $\lambda$ be a measure of welfare gains, and let $\{\iota(\vec{x})\}_{\vec{x}\in X}$ be the partition into indexes implied by $\lambda$. Assume by way of contradiction that there exist functions $\{\mu_{\iota(\vec{x})} : \iota(\vec{x}) \mapsto \mathbb{R}\}_{x\in X}$ and $\{\gamma(\cdot|P) : \{\iota(\vec{x})\}_{\vec{x}\in X} \mapsto \mathbb{R}\}_{P\in D}$ such that each $P \in D^*$ can be represented by $u(\cdot|P) = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$.

Fix some $P \in D^*$ and some $\vec{x} \in X$. Note that there are infinitely many increasing and continuous preference relations that coincide with $P$ on the set $\{\vec{x'} \in \mathbb{R}^J | \vec{x'} \preceq_P \vec{x}\}$. Intuitively, this is because indifference curves outside of this set can be chosen freely (subject to the constraint that they represent increasing preferences). Thus, let $P' \in D^*$ be such that $P' \neq P$ but $P'$ and $P$ coincide on the set $\{\vec{x'} \in \mathbb{R}^J | \vec{x'} \preceq_P \vec{x}\}$.

By assumption, there exists a function $\gamma(\cdot|P')$ such that $P'$ is represented by $\mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P')$. On the set $\{\vec{x'} \in \mathbb{R}^J | \vec{x'} \preceq_P \vec{x}\}$, $P'$ is also represented by $\mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$, and, it thus follows that, for each $\vec{x'}$ such that $\vec{x'} \sim_P \vec{x}$,

$$\mu_{\iota(\vec{x'})}(\vec{x'}) + \gamma(\iota(\vec{x'})|P) = \mu_{\iota(\vec{x})}(\vec{x}) + \gamma(\iota(\vec{x})|P) \Rightarrow \tag{53}$$

$$\gamma(\iota(\vec{x'})|P) - \gamma(\iota(\vec{x})|P) = \mu_{\iota(\vec{x})}(\vec{x}) - \mu_{\iota(\vec{x'})}(\vec{x'})$$

and, similarly

$$\gamma(\iota(\vec{x'})|P') - \gamma(\iota(\vec{x})|P') = \mu_{\iota(\vec{x})}(\vec{x}) - \mu_{\iota(\vec{x'})}(\vec{x'}) \tag{54}$$

Thus,

$$\gamma(\iota(\vec{x'})|P') - \gamma(\iota(\vec{x})|P') = \gamma(\iota(\vec{x'})|P) - \gamma(\iota(\vec{x})|P) \tag{55}$$

Or

$$\gamma(\iota(\vec{x'})|P') = \gamma(\iota(\vec{x'})|P) + (\gamma(\iota(\vec{x})|P') - \gamma(\iota(\vec{x})|P)) \tag{56}$$

Note that, for each $\vec{x''} \in X$, it holds that $e(\vec{x''}, \vec{x}, P) \sim_P \vec{x}$ and $e(\vec{x''}, \vec{x}, P) \in \iota(\vec{x''})$. Consequently, substituting $\vec{x'} = e(\vec{x''}, \vec{x}, P)$ in the above yields the result that, for each $\vec{x''} \in X$, it holds that

$$\gamma(\iota(\vec{x''})|P') = \gamma(\iota(\vec{x''})|P) + c \tag{57}$$

where $c$ is the constant $(\gamma(\iota(\vec{x})|P') - \gamma(\iota(\vec{x})|P))$.

It thus follows that $P'$ is represented by

$$\mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P') = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P) + c \tag{58}$$

which is also a representation of the preferences $P$ (as it is equal to $u(\cdot|P) + c$). It thus follows that $P = P'$, in contradiction to the assumption that $P \neq P'$.

I have thus established that, for any measure of welfare gains, $\lambda$, there cannot exist functions $\{\mu_{\iota(\vec{x})} : \iota(\vec{x}) \mapsto \mathbb{R}\}_{x \in X}$ and $\{\gamma(\cdot|P) : \{\iota(\vec{x})\}_{\vec{x} \in X} \mapsto \mathbb{R}\}_{P \in D}$ such that each $P \in D^*$ can be represented by $u(\cdot|P) = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$. By the second part of Theorem 1, it follows that there does not exist a constitution that jointly satisfies the axioms of Theorem 1.

# G Proof of Corollary 2

Define $\mu_{\iota(\vec{x})}(\vec{x}) = \ln(x_1)$, and $\gamma(\iota(\vec{x})|P) = \ln(h(\vec{x}|P)/x_1)$. Note that $\mu_{\iota(\vec{x})}(\cdot)$ is a function from $\iota(\vec{x})$ to $\mathbb{R}$ that is strictly increasing and continuous, and that $\gamma(\iota(\vec{x})|P)$ is well-defined: if $\iota(\vec{x}) = \iota(\vec{x'})$, then, by Proposition 1, $\vec{x'} = a\vec{x}$ and hence $x'_1 = ax_1$; consequently,

$$\gamma(\iota(\vec{x'}|P) = \ln(h(\vec{x'}|P)/x'_1) = \ln(h(a\vec{x}|P)/(ax_1)) =$$

$$\ln(ah(\vec{x}|P)/(ax_1)) = \ln(h(\vec{x}|P)/x_1) = \gamma(\iota(\vec{x})|P)$$

To see that $u(\cdot|P) = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$ represents the preferences $P$, note that, for each $\vec{x}, \vec{x'} \in X$, it holds that

$$\vec{x} \preceq_P \vec{x'} \Leftrightarrow h(\vec{x}|P) \leq h(\vec{x'}|P) \Leftrightarrow x_1 \frac{h(\vec{x}|P)}{x_1} \leq x'_1 \frac{h(\vec{x'}|P)}{x'_1} \Leftrightarrow$$

$$\ln(x_1 \frac{h(\vec{x}|P)}{x_1}) \leq \ln(x'_1 \frac{h(\vec{x'}|P)}{x'_1}) \Leftrightarrow \ln(x_1) + \ln(\frac{h(\vec{x}|P)}{x_1}) \leq \ln(x'_1) + \ln(\frac{h(\vec{x'}|P)}{x'_1}) \Leftrightarrow$$

$$\mu_{\iota(\vec{x})}(\vec{x}) + \gamma(\iota(\vec{x})|P) \leq \mu_{\iota(\vec{x'})}(\vec{x'}) + \gamma(\iota(\vec{x'})|P) \Leftrightarrow u(\vec{x}|P) \leq u(\vec{x'}|P)$$

Thus, by Theorem 1, the unique constitution that satisfies the axioms is represented by the social welfare function

$$\sum_{i=1}^{I} u(\vec{x_i}|P_i) = \sum_{i=1}^{I}(\mu_{\iota(\vec{x_i}}(\vec{x_i}) + \gamma(\iota(\vec{x_i})|P_i)) = \sum_{i=1}^{I}(\ln(x_{i,1}) + \ln(\frac{h(\vec{x_i}|P_i)}{x_{i,1}})) =$$

$$\sum_{i=1}^{I} \ln(x_{i,1}\frac{h(\vec{x_i}|P_i)}{x_{i,1}}) = \sum_{i=1}^{I} \ln(h(\vec{x_i}|P_i))$$

Thus, the monotone transformation $\exp(\sum_{i=1}^{I} u(\vec{x_i}|P_i))$ represents the social preferences as well, and hence social preferences are represented by

$$\exp(\sum_{i=1}^{I} \ln(h(\vec{x_i}|P_i))) = \prod_{i=1}^{I} \exp(\ln(h(\vec{x_i}|P_i))) = \prod_{i=1}^{I} h(\vec{x_i}|P_i) \tag{59}$$

# H Proof of Corollary 3

As $D$ consists only of homothetic and convex preferences over the set of consumption bundles $C = \mathbb{R}^J_+$, for each $P \in D$, there exists a function $v(\cdot|P) : C \mapsto \mathbb{R}_+$ that is concave, homogeneous of degree 1 and represents $P$'s preferences over $C$. The indirect utility function is then given by

$$u^{ID}(p_1, ..., p_J, m|P) = \max_{c_j} v(c_1, ..., c_J|P) \text{ s.t. } \sum_{j=1}^{J} p_j c_j = m$$

Define $h(m/p_1, ..., m/p_J|P) = u^{ID}(p_1/m, ..., p_J/m, 1|P) = u^{ID}(p_1, ..., p_J, m|P)$. Note that $h$ is increasing, and represents individual preferences over $\tilde{X} = \{(m/p_1, ...., m/p_J)|m, p_j \in \mathbb{R}_+\} = \mathbb{R}^J_+$. To see that it is homogeneous of degree 1, note that, for each $a \in \mathbb{R}_+$,

$$u^{ID}(p_1, ..., p_J, am|P) = \max_{c_j} v(c_1, ..., c_J|P) \text{ s.t. } \sum_{j=1}^{J} p_j c_j = am \tag{60}$$

which can be rewritten as

$$\max_{(c_j/a)} v(a\frac{c_1}{a}, ..., a\frac{c_J}{a}|P) \text{ s.t. } \sum_{j=1}^{J} p_j \frac{c_j}{a} = m$$

Given that $v$ is homogeneous of degree 1, this is the same as

$$\max_{(c_j/a)} av(\frac{c_1}{a},...,\frac{c_J}{a}|P) \text{ s.t. } \sum_{j=1}^{J} p_j \frac{c_j}{a} = m$$

or (changing variables from $c_j/a$ to $c_j$)

$$\max_{c_j} av(c_1,...,c_J|P) \text{ s.t. } \sum_{j=1}^{J} p_j c_j = m$$

Which, by definition, is $au^{ID}(p_1,...,p_J,m|P) = ah(m/p_1,...,m/p_J|P)$, concluding the proof that $h(\cdot|P)$ is homogeneous of degree 1. Thus, individual preferences over $X$ are homothetic.

By Proposition 2, $\lambda_i = ev_i/m_i$ is a measure of welfare gains, which implies the same partition into indexes as in Proposition 1. By Corollary 2, when $D$ is restricted to including homothetic preferences over $\tilde{X}$, there exists a unique constitution that satisfies the axioms of Theorem 1, which can be represented by $\prod_{i=1}^{I} h(m_i/p_{i,1},...,m_i/p_{i,J}|P_i)$. As explained in the text, this social welfare function implies a simple decision rule, according to which the policy change is desirable if and only if $\prod_{i=1}^{I}(1+\lambda_i) > 1$; given that $\lambda_i = ev_i/m$, the corollary follows.

# I Proof of Corollary 4

Define $\mu_{\iota(\vec{x})}(\vec{x}) = f_1(x_1)$ and $\gamma(\iota(\vec{x})|P) = \sum_{j=2}^{J} f_j(x_j|P)$. By assumption, $u(\cdot|P) = \mu_{\iota(\cdot)}(\cdot) + \gamma(\iota(\cdot)|P)$ represents $P$. To see that $\gamma(\iota(\vec{x})|P)$ is well-defined, recall that, by Proposition 3, given this measure of welfare gains, indexes take the form $\iota(\vec{x}) = \{(ax_1, x_2, ..., x_J)|a \in \mathbb{R}_+\}$. Thus, if $\iota(\vec{x}) = \iota(\vec{x'})$, then $\vec{x'} = (ax_1, x_2, ..., x_J)$ for some $a \in \mathbb{R}_+$. Thus, $(x'_2, ..., x'_J) = (x_2, ..., x_J)$ and hence $\gamma(\iota(\vec{x})|P) = \sum_{j=2}^{J} f_j(x_j|P) = \sum_{j=2}^{J} f_j(x'_j|P) = \gamma(\iota(\vec{x'})|P)$.

By Theorem 1, the corollary follows.