Working Paper Series
Department of Economics
University of Verona

Characterizing Inequality Equivalence Criteria

Claudio Zoli

WP Number: 32 November 2012

ISSN: 2036-2919 (paper), 2036-4679 (online)

Characterizing Inequality Equivalence Criteria

Claudio Zoli*
Dipartimento di Scienze Economiche, Università di Verona
.

Abstract

We introduce an axiomatic framework to analyze the perception of inequality across distributions with different total income. The main result is the characterization of a new two parameters generalized version of the inequality equivalence criterion (IEC), the Flexible IEC (FIE). This criterion is able to encompass all the most used criteria of inequality equivalence and is sufficiently flexible to provide a perception consistent with recent evidence from questionnaire investigations, namely that the inequality perception goes from the relative view (focussing on income shares) to the absolute view (focussing on absolute income differentials) as incomes increase. One parameter of the FIE is associated with the Bossert and Pfingsten (Math. Soc. Sc. 1990) Intermediate IEC (IIE) while the other is shown to lead to an alternative single parameter version of the IEC, the Proportional IEC which is dual to the previous. We provide independent characterizations both of PIE and IIE. Also alternative IECs existing in the literature are characterized within the axiomatic framework suggested. These results are consistent with those obtained in the surplus sharing literature by Moulin (Int. J. Game Theory 1987), Young (Math. Op. Res. 1987, J. Econ. Theory 1988) and Pfingsten (Math. Soc. Sc. 1991), and with recent results in inequality measurement by Ebert (Math. Soc. Sc. 2004) and Zheng (Soc. Ch. Wel. 2007). We complete the analysis investigating the implications of extending the domain of the income distributions including negative incomes. The basic consistency properties suggested in order to characterize the IECs are sufficient to select the absolute IEC when unbounded negative incomes are included in the domain.

Keywords: Inequality measurement, inequality equivalence, axiomatic surplus sharing, intermediate inequality.

JEL: D63, D31.

---

*This paper is a revised version of: "A surplus sharing approach to the measurement of inequality". University of York, Discussion Paper 98/25. I am indebted to Walter Bossert, Andrea Brandolini, Vincenzo Denicolo', Udo Ebert, James Foster and Peter Lambert, for helpful comments, discussions and suggestions. Responsibility for any remaining errors is, of course, mine. Mailing address: Dipartimento di Scienze Economiche, Università di Verona, Via dell'artigliere, 19, 37129 Verona, Italy. E-MAIL: claudio.zoli@univr.it

Contents

1 Introduction 3
2 Preliminaries 7
2.1 Axioms for the perception of inequality equivalence 10
3 Results 12
3.1 Path Independence 13
3.2 Betweenness and Flexible Equivalence Criteria 15
3.2.1 Characterizations of Proportional and Intermediate IECs . . . 20
3.3 Connections with the surplus sharing literature 22
3.3.1 Asymmetric distributive rules and further results 26
3.4 Inequality Equivalence Criteria with negative incomes 30
4 Conclusions 32
5 Appendix 33
5.1 Proofs: 33
5.1.1 Proposition 1 33
5.1.2 Proposition 2 37
5.1.3 Proposition 3 41
5.1.4 Proposition 4 44
5.1.5 Proposition 5 45
5.1.6 Proposition 6 48
5.1.7 Proposition 7 55
5.1.8 Proposition 8 57
5.1.9 Proposition 9 58
5.1.10 Proposition 10 58
5.1.11 Proposition 11 61
5.1.12 Proposition 12 62
5.1.13 Proposition 13 63
5.1.14 Proposition 14 63
5.1.15 Proposition 15 67

.

1 Introduction

This paper is about income inequality comparisons between distributions of different total income. We characterize inequality criteria that can be applied to international/intertemporal comparisons and to evaluate distributive effects of alternative policies. These criteria are likely to play a crucial role in the assessment of the relationship between size (or growth) and inequality (or poverty) of income distributions experiencing large differences in average income.

The axiomatic approach to the measurement of income inequality aims at characterizing measures satisfying a set of plausible properties. Some of the axioms suggested (Anonymity and the Principle of Transfers) allow us to make comparisons between distributions of fixed total income and fixed population size, others (the Principle of Population and the Inequality Equivalence Criterion) extend these comparisons over distributions with different population size or different total income[1].

Since the seminal work of Amiel and Cowell (1992), there have been many attempts to test whether individual inequality perceptions are consistent with these axioms. The most controversial axiom appears to be the Inequality Equivalence Criterion (IEC). This criterion specifies the way in which an additional amount of income should be split in order to leave inequality unchanged when compared to the original distribution. Its most used specification (both in theoretical and applied works) is the *relative.* The relative IEC requires that income shares remain unchanged after the addition of a surplus. The main argument in favor of this specification is that it allows the associated inequality rankings to be independent from the units in which income is measured. Kolm (1976) argued that together with the "currency independence" property, the relative criterion also embodies a normative value judgement. Provided we are able to specify a common unit of measure, Kolm (1976) proposed an alternative *absolute* version of the IEC taking into account the differences between the income levels of the individuals instead of their income shares. As shown in Eichhorn and Gehrig (1982) the relative and absolute IECs are not compatible. More precisely, any continuous inequality index satisfying both criteria is constant.[2] In order to capture inequality perceptions that partially combine both approaches, Bossert and Pfingsten (1990) suggests an Intermediate Inequality Equivalence (IIE) criterion which is a parametric representation of an IEC whose extreme cases are the absolute and the relative.

The various IECs suggested in the literature show a common feature: they treat income increments and decrements in the same way irrespective of their size. The idea of following a 'linear expansion' path of inequality equivalence is certainly appealing. However, it is not clear why changes in total income should exhibit the same inequality neutral effect irrespective of whether they come from the distribution of a surplus or from tax payments? Classical findings in the literature on choices under uncertainty (e.g. Kahneman and Tversky, 1978) suggest that individuals facing uncertain prospects tend to evaluate gains and losses in different ways. This seems a reasonable

[1] See, among others, Lambert (2001) and Cowell (2000), and the literature cited therein.

[2] See also Aczel (1987, Sect. 3) and Zheng (1994) where a related result is obtained for poverty measures.

intuitive perspective even for evaluations involving inequality comparisons. For instance, suppose that starting from distribution $A = \{10, 15, 20\}$ the appropriate way to share a surplus of 9 units, in order to leave inequality unchanged, leads to distribution $B_A = \{13, 18, 23\}$ where the surplus is shared according to the absolute criterion. If this is the case, does the appropriate way to collect a revenue from taxation of 27 units lead to the net income distribution $C_A = \{1, 6, 11\}$? What if instead we start from the distribution $A^* = \{10, 10, 100\}$? Is it normatively appealing to consider $C_A^* = \{1, 1, 91\}$ to be inequality equivalent to A^*? In both cases, either going from A to B_A or from A^* to C_A^*, we have respectively added or subtracted a vector where surpluses or tax liabilities have been shared equally. Making use of linear inequality criteria requires that, if we are willing to support the idea that the appropriate way to share a surplus starting from a distribution showing large disparities is to follow the absolute criterion, because in doing so we avoid excessive discrimination of poor individuals as we would have done using a relative criterion, then we must be willing to support the idea that the appropriate way to tax these individuals is trough a "poll tax"! This may contradict the intuition and appear a questionable ethical property for the IEC applied. This point is also at the core of the Kolm-Atkinson early debate on the normative relevance of the relative and absolute inequality equivalence criteria. In support of the absolute inequality view or the less demanding view that inequality should increase because of proportional income increases Kolm (1976a) cites the events of May 1968 in France leading to a 13% increase in all payrolls. As an effect of this decision Kolm (1976a) p. 419 states that: "the Radicals felt bitter and cheated; in their view, this widely increased income inequality". On the other hand Atkinson (1983) provides in favor of the relative inequality view a different historical example considering the problem of a pay cut. The case considered is the one of mutination of sailors of the British navy because of the a shilling a day reduction in the pay of all the naval rating below the rank of a warrant officer in September 1931. Atkinson (1983) p. 6 argues that: "The manifesto of the mutineers made it clear that they were 'quite agreeable to accept a cut that they consider reasonable' [..] but they did not regard it as fair hat they should bear a bigger proportionate cut than the officers."

Both arguments seem to lead the reader to the conclusion that inequality neutral sharing of surpluses and tax liabilities should differ not only (obviously) in sign but also in the distribution of the total amounts. In particular combining Atkinson and Kolm arguments appears reasonable to suggest that surpluses should be shared in more absolute terms while tax liabilities should be shared in more relative terms. This intuition is also consistent with questionnaire evidence showing that neither the relative and absolute IECs[3] nor the parametric IIE criterion[4] are sufficiently flexible to represent the variety of individuals' inequality perceptions. In particular, Amiel and Cowell (1997, 1999) find evidence that "the appropriate inequality equivalence concept depends on the income levels at which inequality comparisons are made". Moreover, they show that as income increases the equivalence concept moves from the relative attitude to the absolute one, a pattern consistent with our intuition. The

[3]See Amiel and Cowell (1992, 1999), Ballano and Ruiz-Castillo (1993), and Harrison and Seidl (1994a,b).

[4]See Ballano and Ruiz-Castillo (1993), and Amiel and Cowell (1997, 1999).

most common IECs applied in the inequality measurement literature are not able to represent this perception since they are all linear.[5]

These issues were also of concern by Sen, in discussing Atkinson index Sen (1973) p. 70 argues that: "What is really restrictive, is [....] the requirement that the inequality measure should be independent of the mean income level. One can argue that for low income levels the inequality measures should take much sharper note of the same degree of *relative* variation on the ground that inequality pinches most when people are closer to starvation. On the other side, I have heard it argued that equality is a 'luxury' that only a rich economy can 'afford', and while I cannot pretend to understand fully this point of view, I am impress by the number of people who seem to be prepared to advocate such a position. Though the considerations run in opposite directions, that in itself is no justification for making (p. 71) the inequality measure *independent* of the level of mean income". In next sentence Sen (1973) p. 71 also highlight a crucial point in the discussion on the appropriate IEC to apply: "We are caught in a bit of a dilemma here. Making inequality measures independent of the mean income seems objectionable, but no alternative general assumption about the relationship of the mean income to these measures seems to be acceptable to all. Also, quantitative specification of the extent of the dependence on the mean income would bring in division even within a camp that may be united on the *direction* of the dependence and *only* on the direction." In this work we provide a solution to the first of the points raised by Sen, identifying a characterization of the IEC that allows to overcome the problem of arbitrariness of the mean income relation of the IECs. The derived IEC is parameterized by two parameters, this functional form will restrict the debate to range of values with the parametric space that can be considered ethically relevant.

We will also provide answers to other questions concerning the structure of the IECs. The only information required by any IECs considered in the literature in order to formalize the way in which a surplus should be distributed to individual i belonging to a group of n persons in order to leave inequality unchanged, is at the most, individual $i's$ income, the amount of the surplus, and the mean (or total) income in the original distribution. It seems therefore surprising that a complicate problem like that of evaluating inequality over distributions of different size could be solved making use of such little information. Do we implicitly disregard some important features of inequality perceptions in making use of criteria relying on such restricted information sets? Our answer to this question will be that as long as the derived IEC has to be combined with the Principle of Transfers the the restriction on the information set is necessary.

We investigate the IEC within an axiomatic framework in order to highlight the normative judgments underlying the selection of an IEC and to identify which properties could characterize equivalence criteria consistent with questionnaires' evidence.[6]

[5]For further IEC see also, Seidl and Pfingsten (1997) and Del Rio and Ruiz Castillo (2000). A criterion showing changes in perception from the relative to the absolute position has been suggested in Krtscha (1994), see also Yoshida (2002).

[6]See Bossert (1998) for a discussion of the relevance of the axiomatic approach compared to the empirical evidence from questionnaires as the appropriate informative basis for policy evaluations.

Instead of imposing axioms directly on the set of income distributions we consider the set of distributive vectors. These vectors are conditioned on a starting distribution, and specify the way in which additional income must be shared in order to leave inequality unchanged with respect to the original distribution.[7]

The analysis is inspired by Moulin (1987) which investigates surplus sharing rules of a joint venture in a cooperative game theoretical framework. Some of the definitions and results can also be traced back to other contributions in the literature concerning sharing rules regarding a division of amounts according to individual liabilities (see Aumann and Maschler (1985), Young (1987, 1988) and Wakker (1988)), or claims (see Chun (1988) and Pfingsten (1991)).

We investigate which properties the distributive rules should satisfy in order to be consistent with the two main axioms in inequality comparisons, Anonymity (A) and the Principle of Transfers (PT). It turns out that consistency with PT requires that the sharing rule applied to divide a surplus can be specified making use only of the information on the surplus shared, the average income of the original distribution and the income of the individual considered. Furthermore we introduce, together with some well known axioms, new ones, we investigate their relationships, and characterize new distributive rules. The key new axiom is called *Betweenness*.[8] It requires that, if the sharing rule of an "infinitesimal" surplus over two starting distributions is the same, then also all distributions obtained as a convex combination of these two must exhibit the same sharing rule.

The innovative aspect of our analysis with respect to the previous surplus sharing literature[9] is that it provides an interpretation of the axioms characterizing the distributive rules in the context of the theory of measurement of income inequality. This approach highlights the suitability of the surplus sharing tools for the purpose of characterizing IECs. Moreover, we provide characterization results (see Proposition 2) involving axioms already applied in the literature whose combined effect has not been yet specified, and present a new characterization of the IIE criterion (see Proposition 9, Part 4). In our opinion, the most relevant aspect is that, making use of the Betweenness property, in Proposition 3 we characterize a two parameters IEC (we call it *Flexible Inequality Equivalence, FIE*) which is sufficiently general to embody the IIE criterion and to represent inequality perceptions moving from positions close to the relative equivalence approaching the absolute equivalence as the average income increases. It is interesting to notice that even if the suggested criterion is more general than the usual ones it is able to characterize only changes of perception going from relative to absolute positions (which are precisely those perceived from questionnaire investigations) and not the reverse.[10]

The IEC associated with the second parameter in the FIE and dual to the IIE (the

[7] Ebert (2004) and Zheng (2002) are recent works providing characterizations of inequality equivalence criteria within a different framework.

[8] Betweenness is conceptually equivalent to the homonymous axiom in the literature on choices under uncertainty, but here it is applied over distributive vectors instead of being applied over distributions as in the original framework.

[9] For a survey see Moulin (2002).

[10] The characterization of the class of inequality indices satisfying FIE and of the Lorenz dominance conditions associated with these indices is given in Zoli (1999).

Proportional IEC) is also characterized in Proposition 8 applying a weakening the requirement of "currency independence" in line with surplus sharing results in Moulin (1987) and recent findings in Zheng (2002) where the implications for inequality measurement of a similar weak requirement are investigated. Further original characterizations of both IIE and PIE are also obtained applying invariance conditions strengthening the Betweenness property. We also present characterization results for the classes of IECs suggested by Seidl and Pfingsten (1997) and Del Rio and Ruiz Castillo (2000) and investigate the implications for our analysis deriving from extending the domain of the income distributions to (unbounded) negative incomes. In this case properties of consistency with PT and A will be sufficient to characterize the absolute criterion as the unique admissible IEC. All formal proofs are to be found in the Appendix.

2 Preliminaries

We assume that the society is composed of n individuals where $n \geq 3$, $n \in \mathbb{N}$ ($\mathbb{N}$ is the set of natural numbers), we will denote by $\mathcal{N} = \{1, 2, 3, ..n\}$ the set of the individuals. Let x_i be the income of individual i. We will denote by $\mathbf{x}$ the $n-$dimensional vector of incomes, by $\mathbf{1}$ [$\mathbf{0}$] the $n-$dimensional vector of ones [zeros], and by $X^n = \left\{\mathbf{x} \in \mathbb{R}^n_+ \backslash \mathbf{0}\right\}$ the set of all feasible income allocations. The average income of a distribution $\mathbf{x} \in X^n$ is $\mu_x = \sum_{i=1}^n x_i/n$. The set of feasible distributions whose total income is μ is $X^n(\mu) := \{\mathbf{x} \in X^n : \mu_x = \mu\}$.

We will denote by $\succcurlyeq$ an (in)equality binary relation over distributions on $X^n \times X^n$, where $\mathbf{x} \succcurlyeq \mathbf{y}$ means that distribution $\mathbf{x}$ is considered at least as equal as distribution $\mathbf{y}$. The reflexive part of this relation is $\sim$, where $\mathbf{x} \sim \mathbf{y}$ means that $\mathbf{x}$ and $\mathbf{y}$ are considered equivalent w.r.t. the concept of equality applied.

We restrict attention to reflexive and transitive inequality binary relations (i.e. orderings) satisfying the following properties:

Definition 1 (Anonymity (A)) $\mathbf{x} \sim \mathbf{P}\mathbf{x}$ *where* $\mathbf{P}$ *is a permutation matrix.*

The anonymity property requires that the only attribute of an individual relevant in inequality comparisons is the income level. Therefore we could consider only the set of ordered vectors $\mathcal{X}^n := \left\{\mathbf{x} \in \mathbb{R}^n_+ \backslash \mathbf{0} : x_i \leq x_j \text{ if } i \leq j\right\}$ denoting with $\mathcal{X}^n(\mu) = X^n(\mu) \cap \mathcal{X}^n$.

Given a distribution $\mathbf{x}$, denote by $\mathbf{x}'$ any distribution obtained from $\mathbf{x}$ applying a *progressive transfer*, that is a *rank-preserving* transfer of income from a rich individual to a poor[11] (hence $\mathbf{x}' \in X^n$ also). We will denote by $\mathcal{P}(\mathbf{x})$ the set of all distributions $\mathbf{x}'$ obtained from $\mathbf{x}$ through a progressive transfer.

Definition 2 (Principle of Transfers (PT)) *Let* $\mathbf{x}, \mathbf{x}' \in X^n$, *if* $\mathbf{x}' \in \mathcal{P}(\mathbf{x})$ *then* $\mathbf{x}' \succcurlyeq \mathbf{x}$.

[11] This definition of progressive transfer and the related Principle of Transfers is due to Fields and Fei (1978).

It is well known that the properties A and PT correspond to imposing S-Concavity on the equality ordering, and that over distributions of same total income and population size the inequality relation is equivalent to the Lorenz ordering[12].

In order to extend the analysis over comparisons involving distributions with different total incomes, we need at least to impose restrictions on $\succcurlyeq_E$considering an *Inequality Equivalence Criterion (IEC)*. An IEC specifies the way in which an additional amount of income (surplus) $\varepsilon > 0$ should be shared in order to leave inequality unchanged with respect to the starting distribution $\mathbf{x}$. Alternatively we could say that the IEC identifies for any feasible income distributions $\mathbf{x}$ a set of feasible distributions with higher total income that are considered inequality equivalent to $\mathbf{x}$.

A parametric IEC encompassing both the relative and the absolute criterion has been formalized by Pfingsten (1986) and Bossert and Pfingsten (1990)[13],

Definition 3 (Intermediate Inequality Equivalence (IIE)) *For all* $\mathbf{x}, \mathbf{y} \in X^n$

$$\begin{array}{lcccl} \textit{if either} & \dfrac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x + \theta} & = & \dfrac{\mathbf{y} - \mu_y \mathbf{1}}{\mu_y + \theta} & \textit{where } \theta \geq 0 \\ \textit{or} & \mathbf{x} - \mu_x \mathbf{1} & = & \mathbf{y} - \mu_y \mathbf{1} & \textit{then} \quad \mathbf{x} \sim \mathbf{y}. \end{array}$$

The IIE criterion can be seen as a relative criterion applied to vectors translated with respect to the origin $-\theta\mathbf{1}$. When $\theta = 0$ the identified criterion is the *relative*, since $\mathbf{x}/\mu_x = \mathbf{y}/\mu_y$. As θ approaches infinity the criterion tends to the *absolute* IEC, $\mathbf{x} = \mathbf{y} + k\mathbf{1}$, where $k = \left(\mu_x - \mu_y\right)$.

Even if this criterion embodies a degree of flexibility through the value of the parameter θ, as shown by Amiel and Cowell (1997, 1999) individuals' inequality perceptions are more general. According to their results as the income increases individuals' perceptions move from relative positions toward the absolute criterion suggesting the existence of a positive relation between θ and μ.

This attitude as well as the IIE are consistent with a widespread accepted general perception (suggested by Dalton (1920) and discussed by Kolm (1976, 1996)) according to which individual perceptions must lie in between the relative and the absolute criteria. This compromise concept of equivalence, in its weakest form, requires that equal proportional increases in all incomes do not decrease inequality, and equal absolute changes act in the opposite direction.

Definition 4 (Compromise Inequality Equivalence (CIE)) *For all* $\mathbf{x}, \mathbf{y} \in X^n$, *such that* $\mu_y \geq \mu_x$

$$\begin{array}{lcll} \mathbf{x} \succcurlyeq \mathbf{y} & \textit{if} & \mathbf{x}/\mu_x = \mathbf{y}/\mu_y \\ \mathbf{y} \succcurlyeq \mathbf{x} & \textit{if} & \mathbf{x} - \mu_x \mathbf{1} = \mathbf{y} - \mu_y \mathbf{1}. \end{array}$$

[12] See Kolm (1969) and Dasgupta, Sen and Starrett (1973).

[13] This formal representation, suggested by Besley and Preston (1991), is obtained from the original one after rearrangements. Ebert (2004) provides a characterization of a general version of the IIE where θ may be negative as long as $-\theta$ is below the lower bound of the income distribution.

In order to accommodate a view consistent with CIE retaining also linearity of the IEC Del Rio and Ruiz Castillo (2000) suggest an intermediate criterion where income increments are shared in a fix proportion of the equal and proportional division. This is a special case of the *Ray Invariant Inequality Equivalence* (RIIE) concept suggested in Seidl and Pfingsten (1997). According to the RIIE criterion, for any ordered distribution, the set of inequality equivalent distributions is on a ray through this distribution which is characterized by a specific ordered vector $\boldsymbol{\upsilon} \in \mathcal{X}^n(1)$ capturing the policy maker view on inequality. That is:

Definition 5 (Ray Invariant Inequality Equivalence (RIIE)) *For all* $\mathbf{x}, \mathbf{y} \in \mathcal{X}^n$ *s.t.* $\mu_y > \mu_x$ *if*

$$\mathbf{y} = \mathbf{x} + \boldsymbol{\upsilon}\left(\mu_y - \mu_x\right)$$

where $\boldsymbol{\upsilon} \in \mathcal{X}^n(1)$ *then* $\mathbf{y} \sim \mathbf{x}$.

In order to avoid controversial results, the set of distributions that are comparable according to this criterion should be bounded conditionally on the vector $\boldsymbol{\upsilon}$ characterizing the equivalence concept. In general the ray invariant distributions are those showing more relative inequality than the vector $\boldsymbol{\upsilon}$.[14] This must be the case because, otherwise, if a distribution shows the same or less relative inequality than $\boldsymbol{\upsilon}$, then this means that it can be obtained from the vector $\boldsymbol{\upsilon}$ adding a distribution which is relatively preferred to the distribution considered. This is not consistent with the idea of CIE, of which the absolute and relative inequality are extreme cases. Moreover, since $\boldsymbol{\upsilon}$ is an ordered vector implies then individuals with the same income in $\mathbf{x}$ may receive a different income in $\mathbf{y}$. Del Rio and Ruiz-Castillo (2000) avoid this problem arguing that any inequality comparison between two distribution $\mathbf{x}, \mathbf{y} \in \mathcal{X}^n$ should not necessarily be independent from the starting distribution $\mathbf{x}$. They adopt a specific RIIE concept according to which the inequality vector $\boldsymbol{\upsilon}$ is characterized as a convex combination (π sharing rule) of the equally distributed vector and the vector of the income shares in the starting distribution, where the parameter of the convex combination is denoted by π.

Definition 6 (Sharing Inequality Equivalence (SIE)) *For all* $\mathbf{x}, \mathbf{y} \in \mathcal{X}^n$, *if*

$$\mathbf{y} = \mathbf{x} + \left(\mu_y - \mu_x\right)\left(\pi \frac{\mathbf{x}}{\mu_x} + (1-\pi)\,\mathbf{1}\right) \tag{1}$$

where $\pi \in [0,1]$ *then* $\mathbf{y} \sim \mathbf{x}$.

This criterion is equivalent to RIIE where $\boldsymbol{\upsilon} = [\pi \frac{\mathbf{x}}{\mu_x} + (1-\pi)\mathbf{1}]$ for all $\mathbf{x} \in \mathcal{X}^n$. The extreme cases of the absolute and relative IECs are obtained when respectively

[14] That is, denoting with $L_x(i/n)$ the Lorenz curve of distribution $\mathbf{x}$ evaluated for the proportion i/n of the population ranked according to the income, for a given $\boldsymbol{\upsilon}$, the set of all possible distributions to which the RIIE could be applied is restricted to the set of all ordered distributions that are Lorenz dominated by $\boldsymbol{\upsilon}$ i.e. the set $\{\mathbf{x} \in \mathcal{X}^n : L_x(i/n) \leq L_v(i/n)\ \forall i \in \mathcal{N}\}$.

$\pi = 0$ and $\pi = 1$. A controversial aspect of SIE is that if $\pi \in (0,1)$ the ray inequality invariant vector may vary as the starting distribution $\mathbf{x}$ moves along the ray of inequality equivalent distributions.[15]

In what follows we will analyze within an axiomatic framework the characterization of IECs. Instead of imposing axioms defined over the set of distributions in X^n, we will characterize the set of distributive vectors connecting inequality equivalent distributions.

2.1 Axioms for the perception of inequality equivalence

Let $\mathbb{R}_{++} := \{\mathbb{R}_+ \backslash 0\}$ be the set of positive real numbers. Denote by $\mathbf{d}(\mathbf{x},\varepsilon) \geq \mathbf{0}$ a function $\mathbf{d}(\mathbf{x},\varepsilon) : X^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}^n_+$ identifying a vector sharing the distribution of the surplus $\varepsilon > 0$ that leaves inequality unchanged w.r.t. distribution $\mathbf{x} \in X^n$, that is

$$\mathbf{x} + \mathbf{d}(\mathbf{x}, \varepsilon) \sim \mathbf{x}.$$

We will call the vector $\mathbf{d}(\mathbf{x},\varepsilon)$ an *Inequality Equivalent Distributive Vector (IEDV).* The element at the i^{th} position of $\mathbf{d}(\mathbf{x},\varepsilon)$ is denoted by $d_i(\mathbf{x},\varepsilon)$. Given the choice of an IEC a unique IEDV for a given $\mathbf{x} \in X^n$ and $\varepsilon > 0$ is identified, and consequently is the restriction on $\succcurlyeq$ imposed by the condition $\mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon) \sim \mathbf{x}$.

We will consider only a particular subset of all possible IEDVs, those where all the surplus is distributed, that is

Axiom 1 (Surplus Clearing (SC)) *For all* $\mathbf{x} \in X^n$, $\varepsilon > 0$, $\sum_{i=1}^n d_i(\mathbf{x},\varepsilon) = \varepsilon$.

According to this property $\mathbf{d}(\mathbf{x},\varepsilon) \in X^n$. We also restrict our analysis to *continuous* IEDVs, i.e. $\mathbf{d}(\mathbf{x},\varepsilon)$ is jointly continuous in ε and in all ranked vectors $\mathbf{x} \in \mathcal{X}^n$.

Axiom 2 (Continuity (C)) *For all* $\mathbf{x} \in \mathcal{X}^n$, $\varepsilon > 0$, *and sequences* $\left\{\mathbf{x}^k,\varepsilon^k\right\}_{k\in\mathbb{N}} \in (\mathcal{X}^n \times \mathbb{R}_{++})$ *if* $\lim_{k\rightarrow\infty}\{\mathbf{x}^k,\varepsilon^k\} = (\mathbf{x}_0,\varepsilon_0)$ *then* $\lim_{k\rightarrow\infty} \mathbf{d}(\mathbf{x}^k,\varepsilon^k) = \mathbf{d}(\mathbf{x}_0,\varepsilon_0)$.

That is we are interested in distributive rules that do not show sudden changes in behavior as we change slightly the surplus to share, or as we apply them to ranked distributions that are very close each others. Moreover we impose that all the IEDVs have piecewise continuous right hand side partial derivative w.r.t. ε. Let $\partial(\mathbf{x},\varepsilon^+) := \lim_{\varepsilon'\rightarrow 0^+} \frac{\mathbf{d}(\mathbf{x},\varepsilon+\varepsilon')-\mathbf{d}(\mathbf{x},\varepsilon)}{\varepsilon'}$.

Axiom 3 (Regularity (R)) *For all* $\mathbf{x} \in X^n$, $\varepsilon > 0$, $\partial(\mathbf{x},\varepsilon^+)$ *exists and is piecewise continuous w.r.t.* ε.

The condition requiring that all $\mathbf{d}(\mathbf{x},\varepsilon)$ admit finite right hand side partial derivative w.r.t. ε for all $\varepsilon > 0$ is imposed as a minimum requirement of regularity of the perception of inequality. We would like to avoid pathological behaviors of $\mathbf{d}(\mathbf{x},\varepsilon)$ that

[15] This does not represent a problem for the analysis in Del Rio and Ruiz-Castillo (2000) because they consider pairwise comparisons of distributions.

we consider of no economic interest, namely we want to exclude inequality perceptions that move in a schizophrenic way as the surplus changes slightly[16]. Moreover, since IEDVs represent the inequality perception across distributions of different total income we impose the minimal requirement that changes in the inequality perceptions due to marginal increases of the surplus are defined.[17]

Considering piecewise continuous functions $\partial(\mathbf{x},\varepsilon^{+})$ allows enough generality to include IECs showing non continuous changes in the way a surplus is shared about thresholds. This is, for instance, the case for IECs whose associated IEDVs share surpluses according to the proportional (relative) rule below a given surplus threshold and share the exceeding part according to the egalitarian (absolute) rule.[18] Given $\mathbf{x} \in X^n$ let

$$\mathcal{D}(\mathbf{x},\varepsilon) := \{\mathbf{d}(\mathbf{x},\varepsilon) \in X^n : \mathbf{d}(\mathbf{x},\varepsilon) \text{ satisfies SC, C, R}\}$$

clearly for all such IEDVs by definition $\mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon) \sim \mathbf{x}$.

We will consider now some extra axioms. These axioms must be satisfied whenever we plan to apply the related IEC together with A or PT in order to characterize inequality measures.

Axiom 4 (Horizontal Equity (HE)) *For all* $\mathbf{x} \in X^n$, $\varepsilon > 0$, *if* $x_i = x_j$ *then* $d_i(\mathbf{x},\varepsilon) = d_j(\mathbf{x},\varepsilon)$.

This is the natural counterpart of the Anonymity axiom in the context of IEDVs. If nothing but income is relevant in inequality evaluations, then the only discrimination between individuals treatment must be in term of their income. Therefore identical individuals must be treated identically. The following axiom specifies in which direction the discrimination moves when incomes differ.

Axiom 5 (Vertical Equity (VE)) *Given* $\mathbf{x} \in \mathcal{X}^n$ *then* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{X}^n$ *for all* $\varepsilon > 0$. *That is, if* $x_i > x_j$ *then* $d_i(\mathbf{x},\varepsilon) \geq d_j(\mathbf{x},\varepsilon)$.

In addition to its intuitive interpretation, VE could be seen as a way of constraining the associated IEC to satisfy the second part of the CIE, together with strict PT. In other words, if equal additions of incomes cannot worsen the inequality then it cannot be that $d_i(\mathbf{x},\varepsilon) < d_j(\mathbf{x},\varepsilon)$ because this is equivalent to applying an equal addition of surplus and a progressive transfer, which by definition will decrease inequality.[19]

The next axiom simply requires consistency between the IEC associated to the characterization of the IEDVs and PT. Denote

$$\mathcal{Y}(\mathbf{x},\varepsilon) := \{\mathbf{y} \in X^n : \mathbf{y} = \mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon),\ \mathbf{x} \in X^n\}.$$

[16] This event is not ruled out by the continuity of $\mathbf{d}(\mathbf{x},\varepsilon)$. It is possible to construct functions that are everywhere continuous in the domain but nowhere differentiable or have nowhere a finite or infinite one-sided derivative. See Gelbaum and Olmsted (1964) ch. 3.8.

[17] Example 2 in Appendix clarifies the meaning of the conditions considered.

[18] It should be point out that R is essential only for some of the results we will present, in many cases it is implied by combinations of the other axioms considered.

[19] It can be shown that for some of the results in next section VE is implied by the other axioms considered, see Remark 4 in the Appendix. Nevertheless we decide to impose it from the beginning. When appropriate we will highlight when it is not necessary for the characterization result.

That is, $\mathcal{Y}(\mathbf{x},\varepsilon)$ is the set of all distributions $\mathbf{y}$ with total income $\varepsilon + \sum_{i=1}^{n} x_i$ that are inequality equivalent to $\mathbf{x}$ according to the IEDVs $\mathbf{d}(\mathbf{x},\varepsilon)$.

Axiom 6 (Lorenz Consistency (LC)) *Consider distributions $\mathbf{x}, \mathbf{x}' \in X^n$, $\mathbf{y}, \mathbf{y}' \in X^n$ such that $\mathbf{y} \in \mathcal{Y}(\mathbf{x},\varepsilon)$, $\mathbf{y}' \in \mathcal{Y}(\mathbf{x}',\varepsilon)$. If $\mathbf{x}' \in \mathcal{P}(\mathbf{x})$, then $\mathbf{y}' \in \mathcal{P}(\mathbf{y})$ where the individuals involved in the progressive transfer are the same, for all $\mathbf{x} \in X^n$, $\mathbf{y} \in \mathcal{Y}(\mathbf{x},\varepsilon)$, and all $\varepsilon > 0$.*

In other words if $\mathbf{x}'$ is obtained from $\mathbf{x}$ through a progressive transfer, and distributions $\mathbf{y}'$ and $\mathbf{y}$ are obtained respectively from $\mathbf{x}'$ and $\mathbf{x}$ sharing the surplus ε, then it must be the case that the surplus sharing procedure is consistent with PT. That is, it has to be that also $\mathbf{y}'$ can be seen as obtained from $\mathbf{y}$ through a progressive transfer occurring between the same individuals involved in the transfer transforming distribution $\mathbf{x}$ into $\mathbf{x}'$.

3 Results

We will first characterize the set of IEDVs $\mathbf{d}(\mathbf{x},\varepsilon)$ satisfying the basic requirements introduced in the previous section, then we will investigate which kind of restrictions on this set are implied by additional axioms. We will try to make explicit the axioms underlying the most used Inequality Equivalence Criteria, and we will compare our results to some available in the surplus sharing literature by Moulin (1987), Chun (1988) and Pfingsten (1991).

This first proposition is closely linked to Theorem 1 in Chun (1988) and Theorem 1 in Pfingsten (1991); it can be obtained from these results by adding VE. Some of the axioms applied for both characterizations are similar (C, SC, HE) others play similar roles but are more general (LC)[20].

Proposition 1 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies conditions HE, LC, VE if and only if there exists a continuous function $\gamma : \mathbb{R}_{++} \times \mathbb{R}_{++} \to [0,1]$ such that for all $\mathbf{x} \in X^n$, and $\varepsilon > 0$*

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\gamma\left(\mu_x,\varepsilon\right)\frac{x_i}{\mu_x} + \left[1-\gamma\left(\mu_x,\varepsilon\right)\right]\right] \quad \forall i \in \mathcal{N}. \tag{2}$$

The logic of the proof is the following. Given two distributions $\mathbf{x}, \mathbf{x}' \in X^n$ where $\mathbf{x}'$ is obtained from $\mathbf{x}$ through a progressive transfer, in order to satisfy LC all the elements of the respective IEDVs associated with the individuals not involved in the transfer should be identical. This avoids to add extra differences in the final vectors not deriving from the original progressive transfer. If we apply an extra progressive

[20] Note that Axiom LC together with SC play the same role as the axiom of *Not Advantageous Reallocation (NAR)* introduced in Moulin (1985a). Although NAR has a strategic relevance in the context of surplus sharing over cooperative ventures, it does not seem to have an immediate intuitive interpretation in the context of economic inequality. In terms of our terminology it requires that the amount of surplus distributed over two individuals involved in a progressive transfer is independent from the distribution of income of all the others, and from the amount of the transfer.

transfer to $\mathbf{x}'$ involving two individuals which are different from those considered in the previous transfer, again all the other individuals must receive the same amount of surplus. For all individuals not involved in both transfers the surplus must be the same in order to be able to rank the final vectors according to PT. Applying a sequence of progressive and regressive transfers keeping fixed the income x_h of only one individual (h), it is possible to obtain the set of all feasible ordered distributions consistent with income x_h at the h^{th} position. For all these distributions, the surplus share of h is unchanged, therefore it depends only on the income x_h on the average (or total) income of the distribution and on the amount of the surplus. Since a progressive transfer changes only the distribution of the incomes of individuals involved without changing the sum, after the transfer the total surplus distributed to the persons not involved remains unchanged. Given SC, also the sum of the surplus of those involved in the transfer is unchanged and is independent from the amount of the transfer.[21] Making use of Th. 1 in Chun (1988) (that considers HE) and adding VE we obtain (2). Even if the result in Proposition 1 is closely linked to those in Chun (1988) and Pfingsten (1991), we prefer to provide extensively the proof in the Appendix, which, we hope, will highlight the connections between the surplus sharing literature and the characterization of inequality equivalence criteria. Most importantly the original intermediate characterization result of IEDVs satisfying LC (in Lemma 1 in Appendix) clarifies that:

Remark 1 *Any IEC that is consistent with the Principle of Transfers in deriving the surplus for individual i must consider only of the information set* $\{x_i, i,\ \mu_x, \varepsilon\}$.

The IEDV in (2) is a convex combination of the relative IEDV $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\frac{\mathbf{x}}{\mu_x}$ and of the absolute IEDV $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{1}$, where the parameter $\gamma(\mu_x, \varepsilon)$ depends respectively on the average income of the distribution at which the evaluation is done and on the amount of the surplus distributed.

We will denote $\mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon) := \{\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x}, \varepsilon)$: $\mathbf{d}(\mathbf{x},\varepsilon)$ satisfies $LC, HE, VE\}$.

3.1 Path Independence

We introduce now an extra axiom discussed in Moulin (1987).

Axiom 7 (Path Independence (PI)) $\mathbf{d}(\mathbf{x}, \varepsilon) + \mathbf{d}[\mathbf{x} + \mathbf{d}(\mathbf{x}, \varepsilon), \varepsilon'] = \mathbf{d}(\mathbf{x}, \varepsilon + \varepsilon')$, $\forall \mathbf{x} \in X^n$, $\forall \varepsilon, \varepsilon' > 0$.

PI requires that the outcome of the surplus sharing procedure is independent from whether the surplus $\varepsilon+\varepsilon'$ is distributed in one go or split into two amounts distributed at different moments. It implies transitivity of the associated IEC.

Proposition 2 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ *satisfies PI if and only if for* $\gamma(\mu,\varepsilon)$ *in* (2) *a continuous (or piecewise continuous) function* $\eta : \mathbb{R}_{++} \to [0, 1]$ *exists such that*

[21] This condition is equivalent to NAR applied in our context, see previous footnote.

$\eta(\mu) := \lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon)$, *and such that for all* $\mathbf{x}\in X^n$, *and* $\varepsilon,\varepsilon' > 0$

$$\mathbf{d}(\mathbf{x},\varepsilon) = (\mathbf{x}-\mu_x\mathbf{1})\left\{\exp[\int_{\mu_x}^{\mu_x+\frac{\varepsilon}{n}}\frac{\eta(\mu)}{\mu}d\mu]-1\right\}+\frac{\varepsilon}{n}\mathbf{1}. \quad (3)$$

It can be proved that for all $\mathbf{d}(\mathbf{x},\varepsilon)\in\mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ PI implies that $\lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon)$ exists and is a continuous (or piecewise continuous) function of μ.[22] Since $\gamma(\mu,\varepsilon)$ is also bounded within $[0,1]$ then also $\eta(\mu) := \lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon)$ is bounded within $[0,1]$. As $\varepsilon\to 0^+$, we obtain the infinitesimal increment in $\mathbf{x}$ that leaves inequality unchanged. Note that the increase in the average income of the distributions along the path identified by $\mathbf{d}(\mathbf{x},\varepsilon)$ is given by the average surplus ε/n. Thus, we can identify each vector as a function of its average μ. The dynamics along the inequality equivalent path is therefore given by the vector $\dot{\mathbf{x}} := \lim_{\varepsilon\to 0^+}\frac{\mathbf{d}(\mathbf{x},\varepsilon)}{\varepsilon/n}$ where $d\mu = \varepsilon/n$. Making use of (2) we can obtain

$$\dot{\mathbf{x}} = \eta(\mu_x)\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x}+\mathbf{1} \quad (4)$$

that allows to identify the set of inequality equivalent distributions as the average income (or equivalently the amount of average surplus to divide) increases. Note that by definition $\mathbf{x}+\mathbf{d}(\mathbf{x},\varepsilon)=\mathbf{y}$ and $d\mu = \mu_y-\mu_x = \frac{\varepsilon}{n}$, it follows that Proposition 2 and Corollary 1 are equivalent, where equation (5) below is obtained as the solution of the above dynamic system.

Corollary 1 *The Inequality Equivalence Criteria associated with the IEDVs characterized in Proposition 2 are: for all* $\mathbf{x},\mathbf{y}\in X^n$, *let* $0<k\leq\min\{\mu_x,\mu_y\}$,[23] *if*

$$\frac{\mathbf{x}-\mu_x\mathbf{1}}{g(\mu_x)} = \frac{\mathbf{y}-\mu_y\mathbf{1}}{g(\mu_y)} \quad \textit{where } g(\mu) := \exp[\int_k^{\mu}\frac{\eta(t)}{t}dt] \quad (5)$$

then $\mathbf{x}\sim\mathbf{y}$.

According to this family of IECs two income vectors are considered inequality equivalent if their normalized distances w.r.t. the mean income are the same, the scale factor for this normalization is the function $g(\mu)$ obtained for values of $\eta(\mu)$ running from 0 to 1.

The IIE can be obtained for $\eta(\mu)=\frac{\mu}{\mu+\theta}$ for all $\mu>0$, whose extreme criteria, the relative and the absolute one are characterized respectively by $\eta(\mu)=1$, and $\eta(\mu)=0$ for all $\mu>0$.

Criteria where discontinuities in $\eta(\mu)$ are identified at some reference levels of the average income are consistent with the characterization in Proposition 2. In the following example we specify from (3) an IEDV associated with a IEC that is relative if applied to all distributions with average income below a threshold $\bar{\mu}>0$ and absolute for all the distributions with average income above the threshold.

[22] See Lemma 3 in the Appendix.

[23] The choice of $k>0$ guarantees that the integrals within square brackets are defined.

Example 1 *Let* $\eta(\mu) = 1$ *if* $\mu < \bar{\mu}$, *and* $\eta(\mu) = 0$ *if* $\mu \geq \bar{\mu}$ *in* (5) *where* $\bar{\mu} > 0$. *The associated IEC is continuous both in* $\mathbf{x}$ *and in* ε *: it states that* $\mathbf{x} \sim \mathbf{y}$ *if*

$$\begin{aligned} \frac{\mathbf{x}}{\mu_x} &= \frac{\mathbf{y}}{\mu_y} && \text{if } \mu_x < \mu_y < \bar{\mu}, \\ \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x} &= \frac{\mathbf{y} - \mu_y \mathbf{1}}{\bar{\mu}} && \text{if } \mu_x < \bar{\mu} \leq \mu_y, \\ \mathbf{x} - \mu_x \mathbf{1} &= \mathbf{y} - \mu_y \mathbf{1} && \text{if } \mu_y > \mu_x \geq \bar{\mu}. \end{aligned}$$

We denote $\mathcal{D}_{\mathcal{P}}(\mathbf{x}, \varepsilon) := \{\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$: $\mathbf{d}(\mathbf{x},\varepsilon)$ satisfies $PI\}$.

3.2 Betweenness and Flexible Equivalence Criteria

We add a new axiom which is in our opinion original in the context of the surplus sharing literature, and has an intuitive interpretation in the context of income inequality measurement. We call it Betweenness because it is logically connected to the homonymous axioms applied in the context of choices under uncertainty. The axiom requires that, given two distributions, *if* it is possible to characterize the IEDVs associated with both distributions obtained sharing an "infinitesimal" surplus, provided that these distributive vectors are the same for both distributions, *then* also for all distributions obtained as a convex combination of the previous, the distributive vectors must exist and are the same.

For a positive ε the vector $\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon$ identifies the shares of the surplus distributed to the agents. For a given starting ordered distribution $\mathbf{x} \in \mathcal{X}^n$ the ordered vector $\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon$ can be interpreted as representing the perception of inequality between distributions with different total income. In general the ordered vector $\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon$ depends on ε, that is it may change according to the difference in total income of the distributions compared. Thus the inequality perception will depend not only on the starting distribution but also on the distance between this one and the other distributions compared. It is therefore difficult to disentangle the effects on inequality perceptions due to each one of these component. Here we suggest to adopt the vector $\boldsymbol{\delta}(\mathbf{x})$ obtained as $\lim_{\varepsilon \to 0^+} [\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon]$ to represent the "local" perception of inequality associated with distribution $\mathbf{x}$ when a positive "infinitesimal" surplus is shared starting from $\mathbf{x}$. The concern for local perceptions is reinforced when path independent IEDVs (or the associated IECs) are considered. As larger amounts of surplus are shared the path identified by the IEDVs connects an infinity of distributions exhibiting increasing average incomes. The inequality perceptions associated with any of these distributions will affect the final vector $\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon$, while the local perception represented by $\boldsymbol{\delta}(\mathbf{x})$ will allow to disentangle the effects of the inequality perception associated with each distribution on the path.

Axiom 8 (Betweenness (B)) *Let* $\boldsymbol{\delta}(\mathbf{x}) := \lim_{\varepsilon \to 0^+} [\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon]$ *for all* $\mathbf{x} \in \mathcal{X}^n$. *Given distributions* $\mathbf{x} \in \mathcal{X}^n(\mu)$; $\hat{\mathbf{x}} \in \mathcal{X}^n(\hat{\mu})$ *where* $\mu \neq \hat{\mu}$, *such that* $\boldsymbol{\delta}(\mathbf{x})$ *and* $\boldsymbol{\delta}(\hat{\mathbf{x}})$ *exist, if* $\boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$ *then* $\boldsymbol{\delta}(\tilde{\mathbf{x}}) = \boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$ *for all* $\tilde{\mathbf{x}}$ *such that* $\tilde{\mathbf{x}} = \beta\mathbf{x} + (1-\beta)\hat{\mathbf{x}}$, $\beta \in [0, 1]$.

The intuition behind property B is the following: (a) only "local" perceptions of inequality are considered in order to eliminate the effects on inequality perceptions due to the distribution of positive surpluses ε connecting distributions with different total incomes. Local perceptions of inequality depend therefore only on the characteristics of the starting distribution considered. (b) These characteristics are not affected by combining two distributions with different total incomes but showing similar local inequality perceptions. For instance suppose that for a given ordered vector of shares $\mathbf{s}$ such that $s_i \geq 0$, $\sum_{i=1}^{n} s_i = 1$, two ordered distributions $\mathbf{x}, \hat{\mathbf{x}} \in \mathcal{X}^n$ obtained as $\mathbf{x} = c_1\mathbf{s} + c_2\frac{\mathbf{1}}{n}$ and $\hat{\mathbf{x}} = \hat{c}_1\mathbf{s} + \hat{c}_2\frac{\mathbf{1}}{n}$, where $c_1, c_2, \hat{c}_1, \hat{c}_2 \geq 0$, $c_1 + c_2 = n\mu$ and $\hat{c}_1 + \hat{c}_2 = n\hat{\mu}$, exhibit the same local inequality perception. Then according to axiom B all distributions $\tilde{\mathbf{x}}$ obtained making use of the set of weights $\tilde{c}_1 = \beta c_1 + (1-\beta)\hat{c}_1$ and $\tilde{c}_2 = \beta c_2 + (1-\beta)\hat{c}_2$ are supposed to exhibit the same local inequality perception. Most of the IECs applied in the literature satisfy this property and in addition also require that distributions showing the same "local" inequality perception are also considered equivalent according to the IEC applied. This is for instance the case for IIE. Here we keep the framework general, allowing for the case where distributions showing the same "local" inequality perceptions are not inequality equivalent.

Property B restricts the set of all possible $\eta(\mu)$ in Proposition 2; it allows to characterize a parametric set of IEDVs, whose associated IECs comprise the IIE.

Proposition 3 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{P}}(\mathbf{x},\varepsilon)$ *satisfies B for all* $\mathbf{x} \in \mathcal{X}^n$*, if and only if either there exist constants* $\lambda \in (0,1]$ *and* $\tilde{\theta} \geq 0$ *such that for all* $\mathbf{x} \in \mathcal{X}^n$

$$\mathbf{d}(\mathbf{x},\varepsilon) = (\mathbf{x} - \mu_x\mathbf{1}) \left\{ \exp[\int_{\mu_x}^{\mu_x+\frac{\varepsilon}{n}} \frac{\lambda}{\mu+\tilde{\theta}} d\mu] - 1 \right\} + \frac{\varepsilon}{n}\mathbf{1}, \tag{6}$$

or $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{1}$.

In other terms, solving (6) ,

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{1} + (\mathbf{x} - \mu_x\mathbf{1}) \left\{ \left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x + \tilde{\theta}}\right)^{\lambda} - 1 \right\} \quad \lambda \in [0,1], \tilde{\theta} \geq 0. \tag{7}$$

Since we consider IEDVs satisfying PI then $\boldsymbol{\delta}(\mathbf{x})$ exists and is unique and according to (4) corresponds to $\frac{1}{n}\left[\eta(\mu)\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x} + \mathbf{1}\right]$. In the proof of Proposition 3 we show that all distributions such that the associated IEDVs belonging to $\mathcal{D}_{\mathcal{P}}(\mathbf{x},\varepsilon)$ satisfy B lie on the rays w.r.t. an appropriate origin $-\tilde{\theta}\mathbf{1} \leq \mathbf{0}$. But differently from what happens for the IIE where $\eta(\mu) = \frac{\mu}{\mu+\tilde{\theta}}$, for the general class of IEDVs derived the parameter θ that characterizes the inequality equivalence perception at each average income level may differ from the one identifying the origin $\tilde{\theta}$, furthermore it may not be fixed. If this is the case then it depends positively on the average income of the distribution.

Like we have done for Proposition 2 and Corollary 1 we can characterize the IEC that is implied by the IEDVs in Proposition 3.

Corollary 2 (Flexible Inequality Equivalence (FIE)) *The IECs associated with each of the IEDVs characterized in Proposition 3 are: for all* $\mathbf{x},\mathbf{y} \in X^n$ *if either*

$$\frac{\mathbf{x}-\mu_x\mathbf{1}}{(\mu_x+\tilde{\theta})^\lambda} = \frac{\mathbf{y}-\mu_y\mathbf{1}}{(\mu_y+\tilde{\theta})^\lambda} \tag{8}$$

where $\lambda \in (0,1]$ *and* $\tilde{\theta} \geq 0$, *or*

$$\mathbf{x}-\mu_x\mathbf{1} = \mathbf{y}-\mu_y\mathbf{1} \tag{9}$$

then $\mathbf{x} \sim \mathbf{y}$.

It is evident that for $\lambda = 1$ the IEC specified is the IIE. If $\tilde{\theta} = 0$, we obtain the IEDV

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{1} + (\mathbf{x}-\mu_x\mathbf{1})\left\{\left(1+\frac{\varepsilon}{n\mu_x}\right)^\lambda - 1\right\} \quad \lambda \in [0,1]\,. \tag{10}$$

This criterion has been characterized by Moulin (1987) (see Theorem 4) in the surplus sharing literature, see §3.3. As a special case of the result in Corollary 2 it is possible to derive a new class of single parameter IECs obtained from FIE when $\tilde{\theta} = 0$. We call this class of IECs: *Proportional Inequality Equivalence.*

Definition 7 (Proportional Inequality Equivalence (PIE)) *For all* $\mathbf{x},\mathbf{y} \in X^n$, *if*

$$\frac{\mathbf{x}-\mu_x\mathbf{1}}{{\mu_x}^\lambda} = \frac{\mathbf{y}-\mu_y\mathbf{1}}{{\mu_y}^\lambda} \quad \textit{where } \lambda \in [0,1]$$

then $\mathbf{x} \sim \mathbf{y}$.

PIE provides representations of the IEC moving from the relative position, when $\lambda = 1$, to the absolute one, when $\lambda = 0$. A specific version of this criterion, valid for $\lambda = \frac{1}{2}$, has been suggested by Krtscha (1994). Recently Yoshida (2002) has also independently generalized Krtscha criterion getting PIE.

The ability to provide a single parameter criterion that is transitive (i.e. satisfies PI) and that moves from the relative to the absolute IECs is an appealing characteristic that is also shared by IIE. These two families of criteria provide a different specifications of the intermediate positions in between the relative and the absolute criteria that furthermore are the only IIE specifications that are consistent with PIE.

Lorenz dominance conditions consistent with Flexible IECs The class of inequality indices that are consistent with Anonymity, Principle of Transfers, and Flexible Inequality Equivalence has been characterized in Zoli (1998, 1999). Following Chakravarty (1988), the unanimous dominance for all these indices applied to distributions in X^n has been shown in Zoli (1999) to be related to a dominance condition based on a parametric transformation of the Lorenz curve expressed in terms of $\tilde{\theta}$ and λ. Letting $L_x(i/n)$ the Lorenz curve of distribution $\mathbf{x} \in X^n$ evaluated at the population percentile i/n, i.e. $L_x(i/n) = \frac{\sum_{j=1}^{i} x_j}{\sum_{j=1}^{n} x_j}$ for $\mathbf{x} \in \mathcal{X}^n$ and for all $i \in \mathbb{N}$, and $L_x(0) = 0$, then the parametric transformation is given by

$$\mathcal{L}_x(k, \lambda, \tilde{\theta}, i/n) := \left(\frac{k+\theta}{\mu_x+\theta}\right)^{\lambda} \mu_x \cdot [i/n - L_x(i/n)]$$

where $k > 0$, and $\tilde{\theta} \geq 0$, $\lambda \in [0, 1]$ are identified by the FIE.

As proved in Proposition 6 in Zoli (1999) distributions $\mathbf{x} \in \mathcal{X}^n$ is (weakly) less unequal than distribution $\mathbf{y} \in \mathcal{X}^n$, for all inequality indices that are consistent with A, PT and FIE (parametrized by $\tilde{\theta}$ and λ) if and only if $\mathcal{L}_x(k, \lambda, \tilde{\theta}, i/n) \geq \mathcal{L}_y(k, \lambda, \tilde{\theta}, i/n)$ for some $k > 0$ and for all $i = 0, 1, 2...n$.

Clearly consistency with PIE requires to compare transformed Lorenz curves obtained for $\tilde{\theta} = 0$ i.e. $\mathcal{L}_x(0, \lambda, 0, i/n) := (\mu_x)^{1-\lambda} \cdot [i/n - L_x(i/n)]$.

Relation with the Intermediate IEC Note that IIE is characterized by either $\eta(\mu) = \frac{\mu}{\mu+\theta}$ or $\eta(\mu) = 0 \; \forall \mu > 0$ in (5), and the IEC in (8) is obtained for $\eta(\mu) = \lambda \frac{\mu}{\mu+\tilde{\theta}}$ $\forall \mu > 0$. If $\lambda \in (0, 1)$ the two criteria do not coincide. At each average income level, when an "infinitesimal" surplus is shared the latter criterion can be considered as a IIE given an appropriate θ. The relation between the two criteria is obtained comparing the respective functions $\eta(\mu)$. For $\lambda \in (0, 1]$, since λ and $\tilde{\theta}$ are constants, we obtain

$$\frac{\mu}{\mu+\theta} = \lambda \frac{\mu}{\mu+\tilde{\theta}} \longleftrightarrow \theta = \left(\frac{1-\lambda}{\lambda}\right)\mu + \frac{\tilde{\theta}}{\lambda}.$$

That is the IEC applied can be interpreted as an Intermediate IEC where the parameter is positively correlated with the average income of the distribution considered. When $\lambda \in (0, 1)$ as the average income increases the intermediate concept associated with the sharing of an "infinitesimal" surplus approaches the absolute inequality criterion, i.e. as income increases the IEDVs approach the equal distribution vector. The above result holds also for the special case of the PIE where $\tilde{\theta} = 0$.

Properties of Flexible IECs In order to interpret the FIE it has to be point out that it is impossible to reconcile the requirement that an IEC satisfies PI and also shares any surplus as a fixed proportion of the relative and the absolute IEDVs. This feature would be a natural appealing characteristic of an IEC satisfying CIE. The IEDVs associated with the SIE satisfy this latter property but are not path independent. If we consider distributions connected using the SIE as their average income increases then income shares are more equalized and a larger proportional component is required to remain on the same path.[24] On the other hand, as pointed out by Seidl and Pfingsten (1997), if we interpret the Intermediate IEC according

[24] Alternatively this pattern can be identified realizing that any IEDV associated with the SIE can be obtained as a IEDV associated with the IIE. For these IEDVs $\mathbf{d}(\mathbf{x}, \varepsilon) = \frac{\varepsilon}{n}(\frac{\mathbf{x}-\mu_x \mathbf{1}}{\mu_x+\tilde{\theta}} + \mathbf{1}) = \frac{\varepsilon}{n}(\frac{\mu_x}{\mu_x+\tilde{\theta}} \frac{\mathbf{x}}{\mu_x} + \frac{\tilde{\theta}}{\mu_x+\tilde{\theta}} \mathbf{1})$ where $\tilde{\theta} \geq 0$, note that the sharing parameter combining the relative and absolute component is $\frac{\mu_x}{\mu_x+\tilde{\theta}}$.

to the proportional procedure, then, as we move from the original distribution $\mathbf{x}$ to another distribution considered inequality equivalent to $\mathbf{x}$ but showing higher average income then the sharing rule approaches the relative division. The reason for this behavior is that the IIE and SIE are linear, therefore once we mix relative and absolute IEDVs the relative inequality content of the obtained distribution is reduced, but according to the linearity of the criterion if we start from the obtained distribution we again have to share a surplus following the same gradient as in the previous step, this requires to lower the weight associated with the absolute part of the gradient.

If we consider any sharing problem as a sequence of problems in which small amounts of surplus are shared in a proportional way, then we get the PIE which is path independent. Note that the *paths identified by FIE (and also by PIE) are non linear*. Furthermore as the average income increases, for sharing of "non-infinitesimal" surpluses it is possible to realize that the distributive vector moves from a relative gradient towards an absolute one. The main difference w.r.t. the IIE is that the "proportional" procedure for IIE is dependent from the average income of the starting distribution.

An intuitive interpretation of the Flexible equivalence concept in (8) is that:

Remark 2 *FIE requires that for each distribution any "infinitesimal" surplus is shared according to a convex combination with parameter $\lambda \in [0,1]$ of the IEDV associated with the IIE whose parameter is $\tilde{\theta} \geq 0$, and the IEDV associated with the absolute IEC.*

Remark 3 *Note that according to PIE (i.e. FIE where $\tilde{\theta} = 0$) any "infinitesimal" surplus is shared as a convex combination, with parameter $\lambda \in [0,1]$, of the relative and absolute rule.*[25]

This result can be showed simply comparing the vectors $\boldsymbol{\delta}(\mathbf{x})$ associated with (8) to those associated with the IIE and the absolute IEC. Denoting by $\tilde{\boldsymbol{\delta}}(\mathbf{x},\lambda,\tilde{\theta}) := \frac{1}{n}\left[\lambda \frac{\mu_x}{\mu_x+\tilde{\theta}} \frac{\mathbf{x}-\mu_x \mathbf{1}}{\mu_x} + \mathbf{1}\right]$ then:

$$\tilde{\boldsymbol{\delta}}(\mathbf{x},\lambda,\tilde{\theta}) = \lambda \cdot \tilde{\boldsymbol{\delta}}(\mathbf{x},1,\tilde{\theta}) + (1-\lambda) \cdot \tilde{\boldsymbol{\delta}}(\mathbf{x},0,\tilde{\theta}). \tag{11}$$

According to this interpretation when $\lambda = 1$, we obtain the IIE with parameter $\tilde{\theta}$; if $\lambda \to 0$, then we approach the absolute inequality limit. As a special case of the PIE case we have, when $\tilde{\theta} = 0$ and $\lambda = \frac{1}{2}$, the criterion suggested by Krtscha (1994) requiring that for every distribution any "infinitesimal" surplus is split according to the unweighted average of the relative and the absolute IEDVs.

Note that for any given μ_x it is always possible to find a constant $\tilde{\lambda} \in [0,1]$ such that $\tilde{\boldsymbol{\delta}}(\mathbf{x},\lambda,\tilde{\theta})$ can be interpreted as a combination of the relative and the absolute IEDVs i.e. $\tilde{\boldsymbol{\delta}}(\mathbf{x},\lambda,\tilde{\theta}) = \tilde{\boldsymbol{\delta}}(\mathbf{x},\tilde{\lambda},0)$, the transformed parameter is $\tilde{\lambda} = \lambda \frac{\mu_x}{\mu_x+\tilde{\theta}}$. This new parameter is not independent from the average income, as $\underset{\sim}{\mu_x}$ increases $\tilde{\lambda}$ increases, and for distributions with sufficiently high average income $\tilde{\lambda}$ converges to λ. The

[25] PIE has been derived by Yoshida (2002) imposing directly this requirement. In our framework this is an implication of the characterization result.

reason is that for very high average incomes the effect of the IIE parameter $\tilde{\theta}$ on the sharing rule applied vanishes and the surplus tend to be shared mainly according to the income shares. When we consider a sharing rule characterized by the coefficient $\tilde{\lambda} = \lambda \frac{\mu_x}{\mu_x + \tilde{\theta}}$ then λ places an upper bound on the value of $\tilde{\lambda}$.

If as a reference criterion for the sharing rule we consider the parameter that identifies the convex combination between the proportional and equal sharing vector of an "infinitesimal" surplus, then for the IEDVs characterized, as incomes increase the sharing rule tends to distribute the surplus according to the proportion λ. On the other hand if the reference criterion adopted to evaluate the sharing rule is the one associated with the value of the IIE parameter θ then as incomes increase θ tends to infinity and the IEDVs are perceived to approach the equal distribution. For discrete changes $\varepsilon > 0$ of the surplus, given path independence the result is implied by the one discussed for marginal surplus changes, as an implication it is possible to show that for a given $\tilde{\theta}$ and μ_x as the surplus to distribute increases the value of θ increases.[26]

3.2.1 Characterizations of Proportional and Intermediate IECs

It seems plausible to require that inequality perceptions satisfy axiom B; this condition is very weak; it implies that a local perception of inequality characterized by a very small change in average income should influence the IEDV in a linear way such that the perception of inequality remains unchanged over linear combinations of two distributions exhibiting the same IEDVs.

In order to narrow the set of IECs then Axiom B could be strengthened in two directions. It could be argued that the Betweenness property should be satisfied for any $\varepsilon > 0$. Alternatively it could be retained the concern for local perceptions but requiring them of being invariant with respect to all combinations of vectors $\mathbf{x}$ and $\hat{\mathbf{x}}$ exhibiting the same local perception and not necessarily to hold only for convex combinations of these vectors.

The first direction of investigation where B is extended to hold all $\varepsilon > 0$ is formalized as follows:

Axiom 9 (Extended Betweenness (EB)) *Let $\mu \neq \hat{\mu}$. Given distributions $\mathbf{x} \in \mathcal{X}^n(\mu)$ and $\hat{\mathbf{x}} \in \mathcal{X}^n(\hat{\mu})$ if $\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\hat{\mathbf{x}},\varepsilon)$ for some $\varepsilon > 0$, then $\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\hat{\mathbf{x}},\varepsilon) = \mathbf{d}(\tilde{\mathbf{x}},\varepsilon)$ for all $\tilde{\mathbf{x}}$ such that $\tilde{\mathbf{x}} = \beta\mathbf{x} + (1-\beta)\hat{\mathbf{x}}$, $\beta \in [0,1]$.*

While the second direction where attention is paid to local perceptions but invariance is required to hold for any combination of the vectors showing the same perception is formalized by:

Axiom 10 (Expanded Local Equivalence (ELE)) *Let $\boldsymbol{\delta}(\mathbf{x}) := \lim_{\varepsilon \to 0^+} [\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon]$ for all $\mathbf{x} \in \mathcal{X}^n$. Given distributions $\mathbf{x} \in \mathcal{X}^n(\mu)$; $\hat{\mathbf{x}} \in \mathcal{X}^n(\hat{\mu})$ where $\mu \neq \hat{\mu}$, such that*

[26] Deriving directly from (10) the value of the IIE parameter θ that makes equivalent, for a given starting distribution, the intermediate IEDV characterized by $\theta \geq 0$ and the general IEDV characterized by $\lambda \in (0,1]$, $\tilde{\theta} \geq 0$ we get that the functional relation identifying θ should satisfy the condition $\frac{\varepsilon}{n}\frac{1}{\mu_x+\theta} = (1 + \frac{\varepsilon}{n}\frac{1}{\mu_x+\tilde{\theta}})^{\lambda} - 1$. That is $\theta = \frac{\varepsilon}{n} \cdot [(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x+\tilde{\theta}})^{\lambda} - 1]^{-1} - \mu_x$ for $\lambda \in (0,1]$, $\tilde{\theta} \geq 0$. From which it can be shown that $\frac{\partial\theta}{\partial\varepsilon} > 0$ and $\frac{\partial\theta}{\partial\mu_x} > 0$ for given $\lambda \in (0,1)$, and $\tilde{\theta} \geq 0$.

$\boldsymbol{\delta}(\mathbf{x})$ and $\boldsymbol{\delta}(\hat{\mathbf{x}})$ exist, if $\boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$ then $\boldsymbol{\delta}(\tilde{\mathbf{x}}) = \boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$ for all $\tilde{\mathbf{x}} \in \mathcal{X}^n$ such that $\tilde{\mathbf{x}} = \alpha\mathbf{x} + \beta\hat{\mathbf{x}}$, where $\alpha > 0, \beta > 0$.

Property EB implicitly requires that all distributions located on the same ray through an origin where all incomes are equal and non-positive exhibit the same IEDV for a given value of the surplus. This behavior is formalized by the following proposition.

Proposition 4 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfies EB if and only if there exist two continuous functions $\omega : \mathbb{R}_{++} \to [0,1]$, and $\theta : \mathbb{R}_{++} \to \mathbb{R}_{+}$ such that for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$:*

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left(\omega(\varepsilon) \cdot \frac{\mathbf{x} - \mu_x\mathbf{1}}{\mu_x + \theta(\varepsilon)} + \mathbf{1}\right). \tag{12}$$

The effect of ε on the sharing rule remains within the brackets, through $\omega(\varepsilon)$ and $\theta(\varepsilon)$, but adding PI imposes that $\omega(\varepsilon) = 1$, and $\theta(\varepsilon) = \tilde{\theta}$ for all $\varepsilon > 0$, which *characterizes the Intermediate* IEC.

Proposition 5 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfies PI and EB if and only if there exist $\tilde{\theta} \geq 0$ such that for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$*

$$\textit{either}\ \ \mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left(\frac{\mathbf{x} - \mu_x\mathbf{1}}{\mu_x + \tilde{\theta}} + \mathbf{1}\right) \tag{13}$$

$$\textit{or}\ \ \mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}. \tag{14}$$

This *new axiomatization of the Intermediate IEC*, highlights the relevance of paying attention to local perceptions of inequality when path independent IECs are considered. In this case a necessary implication of imposing EB is that IECs are linear that is, except for the case of the absolute IEC all distributions exhibiting the same IEDVs for a given $\varepsilon > 0$ turnout also to be inequality equivalent.

If path independency is considered an essential property for the IECs then we can consider ELE that still places attention to local perceptions but requires invariance to hold for all combinations of the vectors exhibiting the same perceptions. This property essentially imposes that local perceptions are invariant for all distributions on the same ray through the origin, i.e. depend on the income shares in each distribution. But it may not be the case that two distributions exhibiting the same local perceptions are also inequality equivalent according to the IEC considered. As a result we obtain a *characterization of the Proportional IEC.*

Proposition 6 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfies PI and ELE if and only if there exist constant $\lambda \in [0,1]$ such that for all $\mathbf{x} \in \mathcal{X}^n$*

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{1} + (\mathbf{x} - \mu_x\mathbf{1})\left\{\left(1 + \frac{\varepsilon}{n\mu_x}\right)^{\lambda} - 1\right\}.$$

To complete the analysis we consider also a further restriction obtained combining EB and ELE thus requiring that the following property is satisfied:

Axiom 11 (Extended Equivalence (EE)) *Let $\mu \neq \hat{\mu}$. Given distributions $\mathbf{x} \in \mathcal{X}^n(\mu)$ and $\hat{\mathbf{x}} \in \mathcal{X}^n(\hat{\mu})$ if $\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\hat{\mathbf{x}},\varepsilon)$ for some $\varepsilon > 0$, then $\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\hat{\mathbf{x}},\varepsilon) = \mathbf{d}(\tilde{\mathbf{x}},\varepsilon)$ for all $\tilde{\mathbf{x}} \in \mathcal{X}^n$ such that $\tilde{\mathbf{x}} = \alpha\mathbf{x} + \beta\hat{\mathbf{x}}$, where $\alpha > 0, \beta > 0$.*

The implications of applying EE can be summarized by a corollary of the previous propositions.

Corollary 3 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfies PI and EE if and only if for all $\mathbf{x} \in \mathcal{X}^n$ either $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}$ or $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\frac{\mathbf{x}}{\mu_x}$.*

Proof. Immediate implication of previous results: EE implies both EB when $\alpha + \beta = 1$ and ELE when we let $\varepsilon \to 0^+$, therefore only the RIE and the AIE result as the intersection of the IIE and the PIE families of IECs. ■

As a result of applying PI and EE we obtain only the relative and the absolute criteria. This result depends on the combination of the properties considered in fact it is possible to show that if PI is dropped we obtain the IEC linked to the IEDV in (12) where $\theta(\varepsilon) = 0$, that is EE implies that $\gamma(\mu_x, \varepsilon)$ in (2) should be independent from μ_x.

3.3 Connections with the surplus sharing literature

In this sections we highlight relationships between previous results and characterizations obtained in the surplus sharing literature, we discuss and reinterpret within our framework results obtained in that literature and suggest further results that can be derived.

In order to investigate these relationships we introduce further axioms. The first axiom requires that the distributive rule applied is independent from the amount of surplus distributed.

Axiom 12 (Linear Equivalence (LE)) *For all $\mathbf{x} \in X^n$, and for all $\varepsilon, \varepsilon' > 0$; $\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon = \mathbf{d}(\mathbf{x},\varepsilon')/\varepsilon'$.*

Since the shares of surplus distributed $\mathbf{d}(\mathbf{x},\varepsilon)/\varepsilon$ are unaffected by the size of the surplus then the IEDVs change proportionally w.r.t. the quantity of surplus and the associated IECs are identified by *linear expansion paths*. Recall that according to (2) $\frac{\mathbf{d}(\mathbf{x},\varepsilon)}{\varepsilon} = \frac{1}{n}\left[\gamma(\mu_x,\varepsilon)\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x} + \mathbf{1}\right]$ for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$. It follows that LE is equivalent to require that $\gamma(\mu,\varepsilon) = \gamma(\mu,\varepsilon')$ for all $\varepsilon, \varepsilon' > 0$.

Next axiom has been discussed in Moulin (1987), Chun (1988) and Pfingsten (1991).

Axiom 13 (Additivity (ADD)) *$\mathbf{d}(\mathbf{x},\varepsilon) + \mathbf{d}(\mathbf{x},\varepsilon') = \mathbf{d}(\mathbf{x},\varepsilon + \varepsilon')$ for all $\mathbf{x} \in X^n$, and for all $\varepsilon, \varepsilon' > 0$.*

ADD is a strong property it induces directly linear equivalence as pointed out in Moulin (1987) Pfingsten (1991), it also implies VE when combined with LC.

Next axiom Weak Currency Independence Consistency (WCIC) can be interpreted as the implication of the weakening of the "Currency Independence" leading to the relative IEC. Instead of imposing $\mathbf{x} \sim \alpha\mathbf{x}$ for all $\alpha > 0$ and all $\mathbf{x} \in X^n$ as "Currency Independence" suggests, we require the inequality ordering to be invariant with respect to changes in the common unit of measure adopted to evaluate the distributions.

Definition 8 (Weak Currency Independence (WCI)) *For all* $\mathbf{x}, \mathbf{y} \in X^n$, $\mathbf{x} \succcurlyeq \mathbf{y} \Leftrightarrow \alpha\mathbf{x} \succcurlyeq \alpha\mathbf{y}$ *for all* $\alpha > 0$.

WCI postulates that changes in the currency adopted to measure income do not affect the inequality comparison between distributions of incomes measured in the same currency. In particular the (as)symmetric part of the binary inequality relation requires that $\mathbf{x}\ (\succ) \sim \mathbf{y} \Leftrightarrow \alpha\mathbf{x}\ \ (\succ) \sim \alpha\mathbf{y}$ for all $\alpha > 0$.[27]

WCI is evidently weaker than "Currency Independence", no prior assumption is imposed on the inequality relation existing between $\mathbf{x}$ and $\alpha\mathbf{x}$ as a result of the change in the currency unit. What is required is that, if distribution $\mathbf{x}$ exhibits the same inequality as distribution $\mathbf{y}$ if measured in Pounds, if we change the currency the new income distributions $\alpha\mathbf{x}$ and $\alpha\mathbf{y}$ measured in Euros should remain inequality equivalent. It may be the case that the inequality associated with $\alpha\mathbf{x}$ is changed with respect to that associated with $\mathbf{x}$, the only restriction is that $\alpha\mathbf{x}$ and $\alpha\mathbf{y}$ have to be considered inequality equivalent.

The WCI requirement of course has also a normative content. If income distributions are measured according to the same unit of measure, then WCI requires that if all incomes are scaled by the same factor then the inequality ranking is not affected, although each distribution may perceive an increase or decrease in inequality.[28]

WCI is a property defined over the inequality binary relation and not in terms of the IEDVs, but it involves restrictions on the set of admissible IECs. Indeed, if we consider $\mathbf{x}$ and $\mathbf{y}$ as obtained through a surplus sharing procedure, as we scale the incomes the equivalence should be retained, this leads to a restriction for the IEDVs. More precisely, consistency with WCI of an IEDV requires $\alpha\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\alpha\mathbf{x},\alpha\varepsilon)$ for all $\mathbf{x} \in X^n$, and $\varepsilon, \alpha > 0$.

Consider distributions $\mathbf{x}, \mathbf{y} \in X^n$ such that $\mathbf{x} \sim \mathbf{y}$, then according to WCI also $\alpha\mathbf{x} \sim \alpha\mathbf{y}$ for all $\alpha > 0$. Suppose that $\mathbf{y} = \mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon)$, then if we scale incomes according to the factor $\alpha > 0$ it must also be that $\alpha\mathbf{y} = \alpha\mathbf{x} + \mathbf{d}(\alpha\mathbf{x},\alpha\varepsilon)$ for all $\mathbf{x} \in X^n$, and $\varepsilon, \alpha > 0$. Note that because of the scaling factor the differences between $\mathbf{y}$ and $\mathbf{x}$ are also scaled by the same factor, hence ε must also be multiplied by α in this second case. Combining both conditions it follows that

$$\alpha\left[\mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon)\right] = \alpha\mathbf{x} + \mathbf{d}(\alpha\mathbf{x},\alpha\varepsilon),$$

[27] Zheng (2002) names "Unit-consistency" an analogous property.

[28] The inequality indices consistent with WCI satisfy the functional equation of homotheticity discussed in Aczél (1987), Candeal and Indurain (1993,1995), Aczél and Moszner (1994) and Bosi, Candeal and Indurain (2000). For applications to welfare measurement consistent with relative inequality indices see Blackorby and Donaldson (1984), Ebert (1987) and Dutta and Esteban (1992). Zheng (2003, 2004) derives decomposable inequality indices and poverty indices satisfying WCIC.

that is:[29]

Axiom 14 (WCI Consistency (WCIC)) $\alpha\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\alpha\mathbf{x},\alpha\varepsilon)$ *for all* $\mathbf{x} \in X^n$, *and all* $\varepsilon, \alpha > 0$.

WCIC is a necessary condition that an IEC has to satisfy if the associated inequality measure is required to satisfy WCI. Of course the conditions imposed on an inequality index by WCI are more restrictive, in particular because they also apply to distribution with the same average income exhibiting the same inequality. It is therefore possible to have indices that satisfy an IEC associated with WCIC but violate WCI.

Normative implications of WCIC. Even if there is agreement on the appropriate unit of measure of income, it is possible to disentangle the *normative implications of WCIC*. Let $\mathbf{s}(\mathbf{x}, \varepsilon) := \frac{\mathbf{d}(\mathbf{x},\varepsilon)}{\varepsilon}$ denote the vector of shares of the IEDV. One could argue that the sharing of the surplus depends on the relative impact of ε compared to the total amount of income in the original distribution. In particular surpluses of the same proportion with respect to the total income of the distribution considered should be shared according to $\mathbf{s}(\mathbf{x}, \varepsilon)$ in the same way if two distributions are identical in terms of income shares. This implies that $\mathbf{s}(\alpha\mathbf{x}, \alpha\varepsilon) = \mathbf{s}(\mathbf{x}, \varepsilon)$ for all $\mathbf{x} \in X^n$, and all $\varepsilon, \alpha > 0$. This is precisely WCIC.

In other words, if the implications of an inequality neutral tax system are considered where $\mathbf{x}$ denotes the post-tax income distribution, ε is the tax revenue and $\mathbf{d}(\mathbf{x}, \varepsilon)$ denotes the vector of tax liabilities, WCIC requires that an inequality neutral tax schedule collecting a fixed proportional revenue should share the tax liabilities in the same way for all the distributions where pre-tax incomes shares are the same. Note that it follows that also post-tax income shares should be the same for the distributions considered, moreover it is not necessary that tax liabilities shares, pre-tax income shares and post-tax income shares have to be the same.

Differently from ADD property WCIC in order to imply VE needs to be paired with LC but also with PI, summarizing:

Remark 4 *For all IEDV's* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x}, \varepsilon)$:
(i) $ADD \Leftrightarrow LE$.
(ii) $LC + ADD \Longrightarrow VE$.
(iii) $LC + PI + WCIC \Longrightarrow VE$.

Characterizations of IIE and PIE. An immediate implication of the Additivity condition of the IEDVs is the characterization of the Intermediate IEC as shown by Pfingsten (1991).

Proposition 7 (Pfingsten (1991)) *An IEC characterized by* $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}(\mathbf{x}, \varepsilon)$ *satisfies LC, HE, PI and ADD if and only if it is the IIE.*

[29] A similar condition has been originally suggested by Moulin (1987) in the surplus sharing literature as "Currency Independence of the surplus".

While following Moulin (1987) in Theorem 4 we can show that WCIC is the key axiom leading to the characterization of the Proportional IEC.

Proposition 8 (Moulin (1987)) *An IEC characterized by* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies LC, HE, PI and WCIC if and only if it is the PIE.*

Notice that VE is not necessary for both results since is implied by LC.

Comparisons with ADD and WCIC help in highlight the connections between these results and those in the previous sections. The following proposition clarifies the connections between the axioms.

Proposition 9 *For any* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$, $\mathbf{x} \in X^n$ *and* $\varepsilon, \varepsilon' > 0$ *the following relations hold:*

(1) $ADD + B \Longrightarrow EB$,

(2) $PI + ADD \Longrightarrow B$,

(3) $PI + ADD \Longleftrightarrow PI + EB$

(4) $PI + WCIC \Longleftrightarrow PI + ELE$

(5) $PI + WCIC + ADD \Longleftrightarrow PI + EE.$

Part 1 is important: it is evident that $ADD \rightarrow B$ for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{P}}(\mathbf{x},\varepsilon)$, because the IIE is a subset of the IEC in (8). What is interesting, and can be realized from Part 2, is that over $\mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ ADD is not sufficient to satisfy B. It is therefore possible to provide a characterization result for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfying ADD and B. From Part 1 we realize that even if B is a more general property than EB, once we add ADD the restrictions imposed by this last axiom are too strong and B+ADD are not sufficiently general as EB for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$. The intuition behind this result is from Remark 4 Part (i). For all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ ADD (as well as LE) rules out any effect of ε on the sharing rule. Parts 3 and 4 trivially derive from comparisons of the results in Propositions 7, 8 and Propositions 5, 6. And Part 5 is a direct implication of the combination of the results in Parts 3 and 4 [see Corollaries 3 and 4].

Moreover, WCIC is stronger than B, as one can expect comparing the characterizations in Propositions 3 and 8. More precisely:

Remark 5 *If* $\delta(\mathbf{x})$ *exists for all* $\mathbf{x} \in \mathcal{X}^n$ *then WCIC implies B.*

Proof. Recall that $\delta(\mathbf{x}) := \lim_{\varepsilon\to 0^+} [d(\mathbf{x},\varepsilon)/\varepsilon]$. According to WCIC $\alpha\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\alpha\mathbf{x},\alpha\varepsilon)$, then $\lim_{\varepsilon\to 0^+} [\alpha d(\mathbf{x},\varepsilon)/\varepsilon] = \lim_{\varepsilon\to 0^+} [d(\alpha\mathbf{x},\alpha\varepsilon)/\varepsilon]$ leading to the condition $\alpha \lim_{\varepsilon\to 0^+} [d(\mathbf{x},\varepsilon)/\varepsilon] = \alpha \lim_{\varepsilon\to 0^+} [d(\alpha\mathbf{x},\alpha\varepsilon)/\alpha\varepsilon]$. It follows that $\delta(\mathbf{x}) = \delta(\alpha\mathbf{x})$ for all $\alpha > 0$, and all $\mathbf{x} \in \mathcal{X}^n$. This condition implies B. ■

Note that as shown in Proposition 9 (Part 2) for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ B is implied by PI+ADD, that is it captures relevant aspects of both PI and ADD without being implied directly by any one of them taken independently. If we strengthen B requiring WCIC then the new property is independent from the combination of PI+ADD and as shown in Corollary 4 selects within the set of IIE criteria just the relative and the absolute criteria.

As already pointed out the relative and absolute IECs are the only Intermediate IECs consistent with the Proportional IEC. It is therefore immediate to conclude that the relative and absolute IEDVs are the only path independent IEDVs satisfying WCIC and ADD.

Corollary 4 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies LC, HE, PI, ADD and WCIC if and only if for all* $\mathbf{x} \in X^n$, *and* $\varepsilon > 0$

$$\text{either } \mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left(\frac{\mathbf{x}}{\mu_x}\right) \text{ or } \mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{1}. \tag{15}$$

This result characterizes the two "extreme" linear IECs satisfying HE. It is consistent with a similar result in Theorem 2 in Moulin (1987). It is also in line with the impossibility result in Proposition 3.3 in Zheng (2002) implicitly pointing out that inequality measures satisfying linear IEC and WCI necessarily have to be relative or absolute. Within our framework an equivalent result could be stated as: properties ADD and WCIC are incompatible with any IEC *strictly* combining relative and absolute views derived from IEDVs satisfying LC, HE, and PI.

3.3.1 Asymmetric distributive rules and further results

We complete the set of characterization results for IECs paying attention to IEDVs that satisfy combinations of WCIC, B, and ADD but may violate either HE (i.e. are non-anonymous or asymmetric) or PI (i.e. are non-transitive). We characterize IEDVs that are combinations of some of those already presented and those associated with the RIIE. Interesting characterizations of SIE are also derived from imposing HE to the IEDVs derived.

Ray invariant equivalence, π sharing rules and proportional equivalence. We provide partial characterizations of the RIIE concept suggested in Seidl and Pfingsten (1997), together with complete characterizations of the "π sharing" equivalence criterion suggested by Del Rio and Ruiz Castillo (2000). The following are partial characterizations of RIIE and liner asymmetric sharing methods.

Proposition 10 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies conditions LC, ADD, and PI if and only if there exist* $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$, *such that* $\upsilon_i \geq \upsilon_{i-1} \geq 0$, $\sum_{i=1}^n \upsilon_i = n$, *for any* $i = 1, 2, ...n$, *and a constant* $\beta \geq 0$ *such that for all* $\mathbf{x} \in \mathcal{X}^n$, $\varepsilon > 0$*: either*

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\frac{\upsilon_i\beta + x_i}{\beta + \mu_x}\right], \quad \forall i \in \mathcal{N}, \tag{16}$$

or

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\upsilon_i, \quad \forall i \in \mathcal{N}. \tag{17}$$

The RIIE is obtained in (17),[30] while (16) provides the asymmetric version of IIE if $\upsilon_i \neq \upsilon_{i-1}$ for some $i \in \mathcal{N}$.[31] Thus, RIIE to the asymmetric equivalent to the

[30] The IEDVs characterized in Proposition 10, have been also characterized in Wakker (1987) in the context of bankruptcy problems making use of an alternative set of axioms.

[31] Note that the IEDVs in (17) and (16) satisfy EB that has been used to characterize the IIE.

absolute IEC specification of the IIE, it has exhibit a greater degree of generality w.r.t. the absolute criterion but at the same time violates an intuitive criterion, which in our opinion has a strong normative justification: it *does not satisfy the HE axiom*.

It is possible to select further the RIIE adding WCIC. If in Corollary 4 we drop HE then the class of IEDVs obtained is expanded including all the "extreme" linear criteria violating HE but satisfying VE. They are the relative IEC and the ray invariant IECs as shown in the following proposition.

Proposition 11 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC, ADD, PI and WCIC if and only if for all $\mathbf{x} \in \mathcal{X}^n$, $\varepsilon > 0$, either*

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\frac{\mathbf{x}}{\mu_x}, \tag{18}$$

or, there exists $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$, such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$, $\sum_{i=1}^n \upsilon_i = n$, for all $i \in \mathcal{N}$, and

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\boldsymbol{\upsilon}. \tag{19}$$

Note that, as already point out the absolute IEC generalizes into the RIIE when HE is not satisfied, while the relative criterion still remains the unique "extreme" linear IEC where income levels are directly considered in the sharing procedure instead of just taking into account the positions in the income ranking. When PI is dropped, then a general version of the SIE of Del Rio and Ruiz-Castillo (2000) that combines it with the RIIE is obtained.

Note that after rearranging (1) SIE requires that for all $\mathbf{x} \in \mathcal{X}^n$ and a given $\pi \in [0,1]$

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left(\pi\frac{\mathbf{x}}{\mu_x} + (1-\pi)\mathbf{1}\right).$$

Proposition 12 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC, ADD and WCIC if and only if there exist a ordered vector $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$, such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$, for all $i \in \mathcal{N}$, $\sum_{i=1}^n \upsilon_i = n$, and a constant $\pi \in [0,1]$ such that for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$,*

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\pi\frac{\mathbf{x}}{\mu_x} + (1-\pi)\boldsymbol{\upsilon}\right]. \tag{20}$$

Rearranging (20) we get $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\pi\mathbf{1} + (1-\pi)\frac{\mathbf{x}}{\mu_x}\right] + \pi\frac{\varepsilon}{n}(\boldsymbol{\upsilon} - \mathbf{1})$ that makes evident that the IEDV can be decomposed into two components, the SIE distributive vector and the difference between the RIIE distributive vector and the equal distribution IEDV. Note that LC requires that the vector $\boldsymbol{\upsilon}$ should be ordered. It should also be pointed out that differently from the SIE a lower value of π does not necessarily implies that the surplus is shared more equally. For instance a feasible vector $\boldsymbol{\upsilon}$ can assign value n only to the richest individual and 0 to all the others. If we let $\hat{\boldsymbol{\upsilon}} = (0,0,0,0...0,n)$ then $d_i(\mathbf{x},\varepsilon) = \pi\varepsilon\left(\frac{x_i}{n\mu_x}\right)$ for all $i \neq n$ and $d_n(\mathbf{x},\varepsilon) = \pi\varepsilon\left(\frac{x_n}{n\mu_x}\right) + (1-\pi)\varepsilon$ i.e. the share $(1-\pi)$ of the surplus is given to the richest individual and the remaining part of the surplus $\pi\varepsilon$ is shared proportionally between the other individuals. In this case a more equal distribution of the surplus is achieved for higher values of π.

Adding HE to the properties considered in Proposition 12 implies that all the elements in $\boldsymbol{\upsilon} \in \mathbb{R}_+^n$ has to be equal, that is $\boldsymbol{\upsilon} = \mathbf{1}$, leading to the following corollary.

Corollary 5 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies conditions LC, HE, ADD and WCIC if and only if there exists a constant* $\pi \in [0,1]$ *such that for all* $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$,

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\pi\frac{\mathbf{x}}{\mu_x} + (1-\pi)\mathbf{1}\right]. \tag{21}$$

This is a characterization of the SIE criterion derived also in Moulin (1987) (see Theorem 3 p. 170).

It is clear that an essential difference between RIIE and SIE is that the former is path independent but does not satisfy HE while the latter violates PI but satisfy HE, both are linear criteria i.e. satisfy ADD (or equivalently LE). It is possible to weaken the previous result of characterization result of SIE. Next proposition could be seen as a corollary of Proposition 4. It may be found surprising the fact that axiom B is involved together with ADD; the relevance of these two axioms is different because the sharing rule does not satisfy PI; in this case B is not implied by ADD.

Proposition 13 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies LC, HE, ADD and B if and only if there exist constants* $\tilde{\omega} \in [0,1]$, *and* $\tilde{\theta} \in \mathbb{R}_+$ *such that for all* $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$:

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left(\tilde{\omega}\cdot\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x+\tilde{\theta}} + \mathbf{1}\right). \tag{22}$$

This result can be obtained considering that ADD implies that $\gamma(\mu,\varepsilon) = \tilde{\gamma}(\mu)$ which is also $\lim_{\varepsilon\to0^+}\gamma(\mu,\varepsilon)$, therefore $\omega(\varepsilon)$ and $\theta(\varepsilon)$ in (12) must be constants. It becomes evident also that for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ $ADD + B \longleftrightarrow ADD + EB$.[32]

The sharing rule in (22) is apparently more general that the one in (21) because it considers a convex combination of the equally distributed vector and the relative vector evaluated w.r.t. the origin $-\tilde{\theta}\mathbf{1}$. In practice, this is only a problem of definition of π. Once we restrict our analysis to a pair of distributions the sharing parameter becomes $\pi = \tilde{\omega}\frac{\mu_x}{\mu_x+\tilde{\theta}}$.

For the sake of completeness it could be proved that adding PI, the condition B is implied by ADD, and $\tilde{\omega}$ is restricted to be equal either to 1 or to 0, which leads to the distributive rule either $d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}(\frac{x_i+\tilde{\theta}}{\mu_x+\tilde{\theta}})$ or $d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}$ for all $i \in \mathcal{N}$ that identifies the Intermediate IEC.

Comparing the result in Corollary5 to the one in Proposition13, it is evident that the only difference between the two characterizations is given by substituting WCIC for B which is weaker, as pointed out in Remark 5. Note that as shown in Proposition 9 (part 2) for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ B is implied by PI+ADD, that is it

[32] Given the result in Remark 4 part (i) the characterization in Proposition 13.can be obtained substituting LE for ADD.

captures relevant aspects of both PI and ADD without being implied directly by any one of them taken independently. If we strengthen B requiring WCIC then the new property is independent from the combination of PI+ADD and as shown in Corollary 4 selects within the set of IIE criteria just the relative and the absolute criteria.

Next proposition generalizes the result in Proposition 8 to IEDVs associated with sharing rules that do not satisfy HE. As one may expect the ordered vector $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$ is substituted to the equal distribution vector $\mathbf{1}$.

Proposition 14 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies conditions LC, PI and WCIC if and only if there exist* $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$ *such that* $\upsilon_i \geq \upsilon_{i-1} \geq 0$, $\sum_{i=1}^n \upsilon_i = n$ *for all* $i \in \mathcal{N}$, $\sum_{i=1}^n \upsilon_i = n$, *and a constant* $\lambda \in [0,1]$ *such that for all* $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$:

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\boldsymbol{\upsilon} + (\mathbf{x} - \mu_x\boldsymbol{\upsilon})\left[\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda} - 1\right]. \tag{23}$$

The IEDVs characterized are a combination of those identifying the RIIE and those leading to PIE, when $\lambda = 0$ we get the RIIE since $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\boldsymbol{\upsilon}$, while when $\lambda = 1$ the proportional IEDV associated with the relative IEC is obtained, $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\boldsymbol{\upsilon} + (\mathbf{x} - \mu_x\boldsymbol{\upsilon})\left[\frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right] = \frac{\varepsilon}{n} \cdot \frac{\mathbf{x}}{\mu_x}$.

According to the result in Proposition 14 the following IEC generalizes PIE when HE is dropped. For all $\mathbf{x}, \mathbf{y} \in \mathcal{X}^n$, $\mathbf{x} \sim \mathbf{y}$ if

$$\frac{x_i - \upsilon_i\mu_x}{(\mu_x)^{\lambda}} = \frac{y_i - \upsilon_i\mu_y}{(\mu_y)^{\lambda}} \quad \text{for all } i \in \mathcal{N}, \tag{24}$$

where $\lambda \in [0,1]$ and $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$ such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$, $\sum_{i=1}^n \upsilon_i = n$.

Note that as anticipated in Remark 4(iii) axiom VE is implied by the other properties. It may also be interesting to consider the benchmark case where $\hat{\boldsymbol{\upsilon}} = (0,0,0,0...0,n)$ then $d_i(\mathbf{x},\varepsilon) = x_i\left(\frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right) - x_i\left[\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right) - \left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda}\right]$ for all $i \neq n$. That is depending on the value of λ the surplus distributed to individual $i \neq n$ is comprised between 0 and the proportional share $x_i\left(\frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)$. While for the richest individual n we have $d_n(\mathbf{x},\varepsilon) = x_n\left(\frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right) + (n\mu_x - x_n)\left[\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right) - \left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda}\right]$, in this case the surplus obtained is comprised between the proportional share and the distribution of all the surplus to the individual. The crucial coefficient is given by $\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda}$.

It may also be useful to recall that when an "infinitesimal" surplus is shared the IEDV becomes

$$\hat{\boldsymbol{\delta}}(\mathbf{x},\lambda,\boldsymbol{\upsilon}) = \frac{1}{n}\left[\lambda\frac{\mathbf{x}}{\mu_x} + \boldsymbol{\upsilon}(1-\lambda)\right] = \lambda \cdot \boldsymbol{\delta}(\mathbf{x},1) + (1-\lambda) \cdot \hat{\boldsymbol{\delta}}(\mathbf{x},0,\boldsymbol{\upsilon}), \text{ for all } \mathbf{x} \in \mathcal{X}^n$$

where $\lambda \in [0,1]$, $\boldsymbol{\delta}(\mathbf{x},1) = \hat{\boldsymbol{\delta}}(\mathbf{x},1,\boldsymbol{\upsilon}) = \frac{1}{n}\left[\frac{\mathbf{x}}{\mu_x}\right]$ and the unit vector $\hat{\boldsymbol{\delta}}(\mathbf{x},0,\boldsymbol{\upsilon}) = \frac{1}{n}\boldsymbol{\upsilon}$ identifies the coordinates of the ordered vector $\boldsymbol{\upsilon}$. Any infinitesimal surpluses is shared as a proportion of the relative rule and the ray invariant vector. For case of vector $\hat{\boldsymbol{\upsilon}}$ we get $\hat{\delta}_i(\mathbf{x},\lambda,\hat{\boldsymbol{\upsilon}}) = \lambda \frac{x_i}{n\mu_x}$ for all $i \neq n$ and $\hat{\delta}_n(\mathbf{x},\lambda,\hat{\boldsymbol{\upsilon}}) = \frac{1}{n}\left[n(1-\lambda) + \lambda \frac{x_n}{\mu_x}\right] = (1-\lambda) + \lambda \frac{x_n}{n\mu_x}$.

Comment The results obtained in this section complete the one in Proposition 8 and those presented in Moulin (1987). They reinforce the relevance of the characterization in Proposition 3. While dropping HE may make available a larger set of possible IECs, it is not clear what are the normative justifications for treating differently individuals in the same economic position. If HE in introduced all the results obtained has to be modified substituting the equal distribution vector $\mathbf{1}$ to $\boldsymbol{\upsilon}$. If we are interested in extending the set of admissible IECs beyond the relative and the absolute criterion then, as shown in Corollary 4, a choice has to be made between PI, ADD and WCIC. In our opinion PI is an essential property for an IEC, while both ADD and WCIC present some interesting aspects but they do not seem essentials. If both properties are applied then we are left only with the relative and absolute criteria, while if only one of them is applied we get either the IIE or the PIE. The Betweenness property combines the essential features of both of them: for all IEDV in $\mathcal{D}(\mathbf{x},\varepsilon)$ satisfying LC, HE, VE and PI we have $ADD \Rightarrow B$ and $WCIC \Rightarrow B$. The fact that Betweenness in addition to its intuitive appealing combines the relevant features of both ADD and WCIC is an additional strong argument in favor of it.

3.4 Inequality Equivalence Criteria with negative incomes

So far we have considered comparisons between non-negative income vectors. This is the most common domain used in inequality analysis, for instance this is the case for all the works considering relative inequality measures[33]. However empirical income distributions may exhibit some negative incomes whose presence may modify some classical results[34]. The results of Hardy, Littlewood and Polya (1934), Berge (1963) and Marshall and Olkin (1979) are derived for larger domains allowing for negative incomes provided the total income is fixed. When distributions of different total incomes are compared Lorenz curves have to be modified in order to accommodate negative incomes, and the relative criterion looses part of its normative interest. For instance the two distributions $(-1,1)$ and $(-10,10)$ are obtained through a relative (scale invariant) transformation but they can hardly be considered inequality equivalent. Moreover, when distributions of utilities are compared requiring that utilities are positive corresponds to introduce a common reference level for their measurement. Considering utility vectors defined over the set of all real numbers then corresponds to eliminating the informational requirement associated with common reference level. The set of admissible IECs will be drastically reduced as the effect of this modification.

Some results presented in related literature (Ebert, 2000, and Ebert and Moyes, 2002) as well as results on characterization of IECs (Ebert, 2001) show that once

[33]See Sen (1973), Eichhorn and Gehrig (1982), Foster (1985), Moyes (1999) and Cowell (2001).

[34]Chen, Tsaur and Rhai (1982) and Berrebi and Silber (1985) present modifications of the Lorenz curve and the Gini index due to the presence of negative incomes.

incomes are allowed to be negative, then the only consistent IEC is the absolute one. These results are derived using a different setting, and the characterization of the IEC is obtained following an approach different from our. We show that if unbounded negative incomes are added to the domain and at least three individuals are considered then the weak requirements of "consistency with the principle of transfers" and "consistency with anonymity" of the IEDVs are sufficient to select the absolute IEC as the unique admissible rule.[35]

We consider IEDVs $\mathbf{d}(\mathbf{x},\varepsilon)$ defined over distributions $\mathbf{x} \in \mathbb{R}^n$, the set of these IEDVs is denoted by $\mathcal{D}_R(\mathbf{x},\varepsilon)$.

The following proposition will clarify the restrictions deriving from imposing LC on the IEDVs in $\mathcal{D}_R(\mathbf{x},\varepsilon)$. The associated IEC does not depend on the income of the individuals but only on the amount of the surplus shared, on the average income of the distribution and on the identity of each individual.

Proposition 15 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_R(\mathbf{x},\varepsilon)$ satisfies condition LC if and only if there exists a sequence of continuous functions $a_i(\mu,\varepsilon) : \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}_+$, satisfying $\varepsilon = \sum_{i=1}^{n} a_i(\mu,\varepsilon)$, such that for all $\mathbf{x} \in \mathbb{R}^n$ and $\varepsilon > 0$:*

$$d_i(\mathbf{x}, \varepsilon) = a_i(\mu,\varepsilon) \qquad \forall i = 1, 2, ..n \tag{25}$$

where $a_i(\mu,\varepsilon) \geq a_j(\mu,\varepsilon)$ if $x_i > x_j$.

Together with the absolute inequality equivalence criterion, also the "ray invariant" criterion suggested by Pfingsten and Seidl (1997) satisfies (25), indeed it is more restrictive, in that it is independent from μ and linear in ε. An essential role for the final characterization is given by the non-negativity of the IEDV. If $d_i(\mathbf{x}, \varepsilon)$ is allowed to be negative, then property LC will require that

$$d_i(\mathbf{x}, \varepsilon) = a_i(\mu,\varepsilon) + b(\mu,\varepsilon)x_i \tag{26}$$

for all $\mu,\varepsilon > 0$, $x_i \in \mathbb{R}$. Once $d_i(\mathbf{x}, \varepsilon) \geq 0$ is imposed then the proportional component $b(\mu,\varepsilon)$ is ruled out. It is reasonable to require that as the total income of the society increases at least the incomes of the poorest individual should not decrease in order to keep inequality unchanged. Otherwise the gap between the poor and the richer individuals would be widened thus leading (at least intuitively) to an increase in inequality.

Note that according to Proposition 15 LC implies VE given that $x_i > x_j$ implies that $a_i(\mu,\varepsilon) \geq a_j(\mu,\varepsilon)$ for all $\mu \in \mathbb{R}$, $\varepsilon > 0$. Once HE is combined with LC and applied to distributions with arbitrary negative incomes, then only the IEDV associated with the absolute criterion survives.

Proposition 16 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_R(\mathbf{x},\varepsilon)$ satisfies conditions LC and HE if and only if for all $\mathbf{x} \in \mathbb{R}^n$ and $\varepsilon > 0$:*

$$d_i(\mathbf{x}, \varepsilon) = \varepsilon/n \qquad \forall i = 1, 2, ..n.$$

[35] If only two individuals are considered then the set of possible IECs is substantially enlarged. For a detailed discussion see Zoli (2002 ch.7).

Proof. HE requires that if $x_i = x_j$ then $d_i(\mathbf{x},\varepsilon) = d_j(\mathbf{x},\varepsilon)$, therefore $a_i(\mu,\varepsilon) = a(\mu,\varepsilon)$ for all i. Substituting into $\varepsilon = \sum_{i=1}^{n} a_i(\mu,\varepsilon)$ we get $a(\mu,\varepsilon) = \varepsilon/n$. ■

This proposition clarifies the tremendous impact on the characterization of IECs associated with the specification of the income domain, in particular when the domain is unbounded from below. Here the absolute equivalence criterion has been characterized adopting a different framework than the one presented in Ebert (2001).

The absolute sharing rule is obtained because of the combination of the basic properties LC and HE and the non-negativity of the elements of the IEDV. A crucial role is played by the assumption that negative incomes can be unbounded. If the domain is modified in order to consider bounded negative incomes, for instance setting $x > x_0$ then plenty of IECs can be derived.

For instance the Flexible Inequality Equivalence criterion can be specified in order to be applied to income distributions where all $x > x_0$ and $x_0 < 0$ setting $\tilde{\theta} \geq -x_0$. The IEC obtained will require that $\mathbf{x}$ and $\mathbf{y}$ exhibit the same inequality if for all individuals $i = 1, 2, ...n$

$$\frac{y_i - \mu_y}{(\mu_y + \tilde{\theta})^{\lambda}} = \frac{x_i - \mu_x}{(\mu_x + \tilde{\theta})^{\lambda}} \quad \text{where } 1 \geq \lambda \geq 0, \tilde{\theta} \geq -x_0 \tag{27}$$

The associated IEDVs are as in (7) where $\tilde{\theta} \geq -x_0$. They are non-negative and satisfy LC, HE and VE as well as PI. For $\lambda = 1$ the IEDVs satisfy also ADD. They are $\mathbf{d}(\mathbf{x}, \varepsilon) = \frac{\varepsilon}{n}\left[\frac{\mathbf{x}+\tilde{\theta}\mathbf{1}}{\mu_x+\tilde{\theta}}\right]$ for $\tilde{\theta} \geq -x_0$ and are associated with the extension of Bossert and Pfingsten's (1990) intermediate inequality equivalence criterion characterized in Ebert (2001) when income distributions with bounded negative incomes are considered.

4 Conclusions

We discuss an axiomatic framework for the analysis of IECs. A general characterization of a new two parameters, flexible IEC is provided, the *Flexible Inequality Equivalence* (FIE). The FIE is consistent with questionnaire evidence on the perception of inequality over distributions with different total income suggesting that attitudes toward inequality move from relative to absolute positions as incomes increase. FIE is a two parameters IEC. If the effect of one parameter is eliminated we get the *Intermediate Inequality Equivalence* (IIE) while if the effect of the other parameter associated with the IIE is eliminated we get a new IEC the *Proportional Inequality Equivalence* (PIE). This criterion provides a formalization of inequality attitudes included in between the relative and the absolute that is alternative to that suggested by the IIE. A complete characterization of PIE is also presented requiring consistency with a weak version of the "currency independence" property.

New characterizations of known IECs are also provided including IIE, Ray Invariant Inequality Equivalence and Sharing Inequality Equivalence. The implications deriving from combinations of the suggested axioms are explored and a variety of additional IECs are characterized and analyzed in order to complete the analysis. We believe that the approach proposed, is helpful in analyzing the different merits of various IECs. In particular it allows to highlight the normative relevance of *non-linear*

IECs, like FIE (and its special case PIE) that have been neglected in the literature on inequality measurement.

It is important to recall that except for the PIE and the IECs satisfying "weak currency independence consistency" property, all the discussed IECs can be applied only if there is consensus on the unit in which income is measured.

5 Appendix

Example highlighting the relevance of the Regularity axiom.

Example 2 *The following IEDV*

$$\mathbf{d}(\mathbf{x},\varepsilon) = \begin{cases} \frac{\varepsilon}{n\mu_x}\mathbf{x} & \text{if } 0 < \varepsilon \leq \bar{\varepsilon} \\ \frac{(\varepsilon+\bar{\varepsilon})}{2}\frac{\mathbf{x}}{n\mu_x} + \frac{(\varepsilon-\bar{\varepsilon})}{2}\left[\frac{\mathbf{1}}{n} + \sin\left(\frac{1}{\varepsilon-\bar{\varepsilon}}\right)\frac{\mu_x\mathbf{1}-\mathbf{x}}{n\mu_x}\right] & \text{if } \varepsilon > \bar{\varepsilon} \end{cases}$$

satisfies both SC and C. For $\varepsilon > \bar{\varepsilon}$, the marginal impact of an increase in surplus is

$$\begin{aligned} \partial(\mathbf{x},\varepsilon^{+}) &= \frac{1}{2}\frac{\mathbf{x}}{n\mu_x} + \frac{1}{2}\left[\frac{\mathbf{1}}{n} + \sin\left(\frac{1}{\varepsilon-\bar{\varepsilon}}\right)\frac{\mu_x\mathbf{1}-\mathbf{x}}{n\mu_x}\right] - \frac{1}{2(\varepsilon-\bar{\varepsilon})}\left[\cos\left(\frac{1}{\varepsilon-\bar{\varepsilon}}\right)\frac{\mu_x\mathbf{1}-\mathbf{x}}{n\mu_x}\right] \\ &= \frac{1}{2}\frac{\mathbf{x}+\mu_x\mathbf{1}}{n\mu_x} + \frac{1}{2}\left(\frac{\mu_x\mathbf{1}-\mathbf{x}}{n\mu_x}\right)\left[\sin\left(\frac{1}{\varepsilon-\bar{\varepsilon}}\right) - \frac{1}{(\varepsilon-\bar{\varepsilon})}\cos\left(\frac{1}{\varepsilon-\bar{\varepsilon}}\right)\right] \end{aligned}$$

As ε approaches $\bar{\varepsilon}$ then $\partial(\mathbf{x},\varepsilon^{+})$ is not defined given that (i) both the limits of sin and cos functions are not defined for $\varepsilon \to \bar{\varepsilon}$, and (ii) $\frac{1}{(\varepsilon-\bar{\varepsilon})}$ is unbounded. These two types of situations are avoided imposing R. Note that as $\varepsilon \to \bar{\varepsilon}$ the period of $\sin\left(\frac{1}{\varepsilon-\bar{\varepsilon}}\right)$ tends to 0, and the frequency of the jumps between the values 1 and -1 tends to infinite: the marginal increase in the surplus $(\varepsilon - \bar{\varepsilon})$ is shared in a proportional way (for $\sin = -1$) and in an egalitarian way (for $\sin = 1$) with "infinite" frequency.

5.1 Proofs:

5.1.1 Proposition 1

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies conditions HE, LC, VE if and only if there exists a continuous function $\gamma : \mathbb{R}_{++} \times \mathbb{R}_{++} \to [0,1]$ such that for all $\mathbf{x} \in X^n$, and $\varepsilon > 0$

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\gamma(\mu_x,\varepsilon)\frac{x_i}{\mu_x} + [1-\gamma(\mu_x,\varepsilon)]\right] \quad \forall i \in \mathcal{N}. \tag{28}$$

Proof of Proposition 1:

Sufficiency: Check that (28) satisfies conditions VE, HE, LC.

Necessity: In order to provide the necessity part we will make use of the following result:

Lemma 1 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC if and only if there exists a set of continuous functions $a_i(\mu,\varepsilon) : \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}_{+}$ for all $i \in \mathcal{N}$ satisfying (i) $a_j(\mu,\varepsilon) \geq$*

$a_i(\mu,\varepsilon) \geq 0$ for all $\mu,\varepsilon > 0$, if $x_j \geq x_i$, (ii) $\varepsilon + \min_i \{a_i(\mu,\varepsilon) \cdot (n - i + 1) : i \in \mathcal{N}\} \geq \sum_{h=1}^{n} a_h(\mu,\varepsilon)$ for all $\mu,\varepsilon > 0$, where $x_{i+1} \geq x_i$, and (iii) $\varepsilon + n\mu \geq \sum_{h=1}^{n} a_h(\mu,\varepsilon)$, for all $\mu,\varepsilon > 0$, such that for all $\mathbf{x} \in X^n$ and all $\varepsilon > 0$:

$$d_i(\mathbf{x}, \varepsilon) = a_i(\mu,\varepsilon) + \frac{\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon)}{n\mu} x_i \qquad \forall i \in \mathcal{N}. \tag{29}$$

Proof of Lemma 1:

Consider two distributions $\mathbf{x}, \mathbf{x}' \in X^n$ where $\mathbf{x}'$ is obtained from $\mathbf{x}$ through a progressive transfer of the amount δ from the individual at position j to the individual at position $i < j$. Denoting by $\boldsymbol{\delta}_{j,i}$ the vector $(0_1, 0_2, 0_{i-1}, +\delta_i, 0_{i+1}, ..0_{j-1}, -\delta_j, 0_{j+1}, ..0_n)^T$, then $\mathbf{x}' = \mathbf{x} + \mathbf{P}_{\mathbf{x}}^T \boldsymbol{\delta}_{j,i}$, where $\mathbf{P}_{\mathbf{x}}$ is the permutation matrix which ranks incomes in $\mathbf{x}$ in increasing order, thus $\mathbf{P}_{\mathbf{x}}\mathbf{x} \in \mathcal{X}^n$. Let $\rho(i)$ denote the rank on $\mathbf{P}_{\mathbf{x}}\mathbf{x}$ of individual i in the original distribution $\mathbf{x}$, therefore $\rho^{-1}(j) = j'$ identifies the individual in the original distribution to whom is associated the j^{th} position in the ordered distribution. In order to satisfy condition LC we need that $d_k(\mathbf{x},\varepsilon) = d_k(\mathbf{x}',\varepsilon)\ \forall k \in \mathcal{N}, k \neq i', j'$; and $d_{i'}(\mathbf{x}',\varepsilon) \geq d_{i'}(\mathbf{x},\varepsilon) - \delta$, $d_{j'}(\mathbf{x}',\varepsilon) \leq d_{j'}(\mathbf{x},\varepsilon) + \delta$. It is also necessary that if $x_{j'} > x_{i'}$ then also $x_{j'} + d_{j'}(\mathbf{x},\varepsilon) > x_{i'} + d_{i'}(\mathbf{x},\varepsilon)$ for all $i', j' \in \mathcal{N}$, all $\mathbf{x} \in X^n$, and all $\varepsilon > 0$. That is, the relative position in the income ranking of $i', j' \in \mathcal{N}$ has to be preserved after the IEDV is added. Because this condition is satisfied for all $\mathbf{x} \in X^n$ then it is implicitly satisfied also $x_{j'} - \delta + d_{j'}(\mathbf{x}',\varepsilon) > x_{i'} + \delta + d_{i'}(\mathbf{x}',\varepsilon)$ if $x_{j'} - \delta > x_{i'} + \delta$.

Apply now a regressive transfer $\boldsymbol{\delta}_{l,m}$ where $l < m$, to the distribution $\mathbf{x}'$, we obtain $\mathbf{x}'' = \mathbf{x}' + \mathbf{P}_{\mathbf{x}}^{\mathbf{T}} \boldsymbol{\delta}_{l,m}$. The result of this procedure is exactly the same as that obtainable applying a progressive transfer $\boldsymbol{\delta}_{m,l}$ to $\mathbf{x}''$; this implies together with LC that $d_k(\mathbf{x}'',\varepsilon) = d_k(\mathbf{x}',\varepsilon)\ \forall k \in \mathcal{N}, k \neq l', m'$. Fix now the income (x_h) of individual h; applying a sequence of progressive and regressive transfers of the kind described above we can obtain all possible distributions belonging to $\mathcal{X}^n(\mu)$ such that $x_{\rho(h)}$ is fixed, $x_{j-1} \leq x_j \leq x_{\rho(h)}\ \forall j = 2, 3, ..n-1$, where $\rho(h) > j \geq 2$; and $x_{\rho(h)} \leq x_{i-1} \leq x_i \leq x_n\ \forall i = 3, 4, ..n$, where $\rho(h) < i \leq n$. Denote with $\mathcal{X}_h^n(\mu)$ the set of all such distributions. Condition LC implies that $d_h(\mathbf{x},\varepsilon) = d_h(\mathbf{x}',\varepsilon)$ for all $\mathbf{x}, \mathbf{x}' \in X^n$ such that $\mathbf{P}_{\mathbf{x}}\mathbf{x}, \mathbf{P}_{\mathbf{x}}\mathbf{x}' \in \mathcal{X}_h^n(\mu)$, for all $h \in \mathcal{N}$, for all x_h such that $\mathbf{x}, \mathbf{x}' \in X^n(\mu)$, and for all $\mu > 0$. This implies that $d_h(\mathbf{x},\varepsilon)$ does not depend on the distribution of the incomes but only on the level of income of the h^{th} individual, possibly on his/her position $\rho(h)$, on the average income, and obviously on ε, for all $h \in \mathcal{N}$. That is

$$d_h(\mathbf{x},\varepsilon) = f_h(x_h, \mu,\varepsilon),$$

where the function $f_h(.)$, which is continuous, implicitly takes into account $\rho(h)$, therefore could be different for each individual, and $\sum_h f_h(.) = \varepsilon$. According to this representation, given LC, we have $f_i(x_i + \delta, \mu,\varepsilon) \geq f_i(x_i, \mu,\varepsilon) - \delta$, and $f_j(x_j - \delta, \mu,\varepsilon) \leq f_j(x_j, \mu,\varepsilon) + \delta$. Recalling the condition $\sum_h f_h(.) = \varepsilon$ we obtain[36]:

$$f_i(x_i + \delta, \mu,\varepsilon) + f_j(x_j - \delta, \mu,\varepsilon) = f_j(x_j, \mu,\varepsilon) + f_i(x_i, \mu,\varepsilon)$$

for any $i, j \in \mathcal{N}$; $x_i, x_j \geq 0$ such that $x_j > x_i$ and $\delta > 0$ chosen in a way such that the transfer does not alter the position of the individuals involved. Denoting with

[36] This is the condition of Non Advantageous Reallocation suggested in Moulin (1987).

$g_i(x_i, \delta, \mu,\varepsilon) := f_i(x_i + \delta, \mu,\varepsilon) - f_i(x_i, \mu,\varepsilon)$, then

$$g_i(x_i, \delta, \mu,\varepsilon) = g_j(x_j - \delta, \delta, \mu,\varepsilon)$$

for any $i, j \in \mathcal{N}$; for x_i, x_j such that $0 \leq x_i < x_j \leq n\mu$ and $\delta > 0$ chosen in a way such that the transfer does not alter the position of the individuals involved.

This implies that $g_i(x_i, \delta, \mu,\varepsilon) = g(\delta, \mu,\varepsilon)$, that is $g_i(x_i, \delta, \mu,\varepsilon)$ does not depend on i and x_i for all $i \in \mathcal{N}$, $x_i \geq 0$. Suppose that this is not the case, consider what happens if $x_i = \alpha \geq 0$ and $x_j = \alpha + \delta$,[37] then $g_i(\alpha, \delta, \mu,\varepsilon) = g_j(\alpha, \delta, \mu,\varepsilon)$ for every $\alpha \geq 0$, which implies that $g_i(\alpha, \delta, \mu,\varepsilon) = g(\alpha, \delta, \mu,\varepsilon)$. If this is the case then $g(x_i, \delta, \mu,\varepsilon) = g(x_j - \delta, \delta, \mu,\varepsilon)$. Letting $x_i = \alpha$ and $x_j = \beta + \delta$, then $g(\alpha, \delta, \mu,\varepsilon) = g(\beta, \delta, \mu,\varepsilon)$ for any $\alpha, \beta \geq 0$ which makes evident that $g(x_i, \delta, \mu,\varepsilon)$ does not depend on x_i. Thus,

$$f_i(x_i, \mu,\varepsilon) + g(\delta, \mu,\varepsilon) = f_i(x_i + \delta, \mu,\varepsilon)$$

for any $\delta > 0$, and for all $\mu,\varepsilon > 0$, $x_i \geq 0, i \in \mathcal{N}$, where $g(0, \mu,\varepsilon) = 0$. The solution of this functional equation is obtained as

$$f_i(x_i, \mu,\varepsilon) = a_i(\mu,\varepsilon) + b(\mu,\varepsilon)x_i \quad \text{for all } \mu,\varepsilon > 0,\ x_i \geq 0,\ i \in \mathcal{N},$$

see Aczél (1966, Th. 1 p.142 and th.1 p.34, see also p.85) and Eichhorn (1978) Ch.1 for discussion of domain restrictions.

Consider condition $\sum_h f_h(.) = \varepsilon$. Specifying $\sum_i [a_i(\mu,\varepsilon) + b(\mu,\varepsilon)x_i] = \varepsilon$, that is

$$\sum_{i=1}^{n} a_i(\mu,\varepsilon) + b(\mu,\varepsilon)n\mu = \varepsilon \leftrightarrow b(\mu,\varepsilon) = \frac{\varepsilon - \sum_{i=1}^{n} a_i(\mu,\varepsilon)}{n\mu},$$

thus:

$$d_i(\mathbf{x},\varepsilon) = f_i(x_i, \mu,\varepsilon) = a_i(\mu,\varepsilon) + \frac{\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon)}{n\mu} x_i \quad \forall i \in \mathcal{N}. \tag{30}$$

Consider the condition $f_i(x_i + \delta, \mu,\varepsilon) \geq f_i(x_i, \mu,\varepsilon) - \delta$, which could be rewritten as:

$$a_i(\mu,\varepsilon) + \frac{\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon)}{n\mu} (x_i + \delta) \geq a_i(\mu,\varepsilon) + \frac{\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon)}{n\mu} x_i - \delta \quad \forall i \in \mathcal{N}, \tag{31}$$

that is

$$\frac{\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon)}{n\mu} + 1 \geq 0 \quad \forall i \in \mathcal{N}, \tag{32}$$

which is satisfied if and only if $\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon) + n\mu \geq 0$.

If we let $\mathbf{x} = (0, 0, ..., n\mu)$, in order to satisfy $d_i(\mathbf{x},\varepsilon) \geq 0$ for all $i = 1, 2..n$ where $x_n \geq x_i$ for all $i \in \mathcal{N}$ it must be that $a_i(\mu,\varepsilon) \geq 0$ for all $i = 1, 2..n - 1$ and $\varepsilon \geq \sum_{i=1}^{n-1} a_i(\mu,\varepsilon) \geq 0$ for all $\mu,\varepsilon > 0$. Finally the condition requiring that if $x_j > x_i$

[37] Notice that it is not necessary that x_i, x_j belong to the same distribution; the functions $f_j(x_j, \mu,\varepsilon)$ apply to all distributions belonging to $\mathcal{X}^n(\mu)$ for any income x_j, position j and surplus ε, otherwise this example is not consistent with the condition that δ is chosen in a way such that the transfer does not alter the position of the individuals involved.

then also $x_j + d_j(\mathbf{x},\varepsilon) > x_i + d_i(\mathbf{x},\varepsilon)$ for all $i, j \in \mathcal{N}$, all $\mathbf{x} \in X^n$, and all $\varepsilon > 0$ can be restated as

$$a_j(\mu,\varepsilon) + \left[\frac{\varepsilon - \sum_{j=1}^n a_j(\mu,\varepsilon)}{n\mu} + 1\right] x_j > a_i(\mu,\varepsilon) + \left[\frac{\varepsilon - \sum_{j=1}^n a_j(\mu,\varepsilon)}{n\mu} + 1\right] x_i \quad (33)$$

for all $i, j \in \mathcal{N}$, $\varepsilon, \mu > 0$ if $x_j > x_i \geq 0$. Let $x_j - x_i = k_{ij}$ where $n\mu \geq k_{ij} > 0$, rearranging the previous condition we get

$$a_j(\mu,\varepsilon) - a_i(\mu,\varepsilon) > -\left[\frac{\varepsilon - \sum_{j=1}^n a_j(\mu,\varepsilon)}{n\mu} + 1\right] k_{ij} \quad \forall i, j \in \mathcal{N}, \varepsilon, \mu, k_{ij} > 0 \quad (34)$$

where the term into square brackets is non-negative given (32) . Taking the limit for $k_{ij} \to 0^+$ we obtain $a_j(\mu,\varepsilon) - a_i(\mu,\varepsilon) \geq 0 > \lim_{k_{ij} \to 0^+} -\left[\varepsilon - \sum_{j=1}^n a_j(\mu,\varepsilon) + n\mu\right] k_{ij}/n\mu$. Thus $a_j(\mu,\varepsilon) \geq a_i(\mu,\varepsilon)$ for all $i, j \in \mathcal{N}$, $\varepsilon, \mu > 0$ if $x_j > x_i \geq 0$. This condition implies also that $a_n(\mu,\varepsilon) \geq 0$ given that $a_i(\mu,\varepsilon) \geq 0$ for all $i = 1, 2, ..n - 1$.

From considering the distribution $\mathbf{x} = (0, 0, ..., n\mu)$, and imposing $d_i(\mathbf{x},\varepsilon) \geq 0$ for all $i \in \mathcal{N}$ we derived the condition $\varepsilon + a_n(\mu,\varepsilon) \geq \sum_{h=1}^n a_h(\mu,\varepsilon) \geq 0$ for all $\mu,\varepsilon > 0$. Knowing that $a_i(\mu,\varepsilon) \geq 0$ for all $i \in \mathcal{N}$ we can derive the more general condition implying $d_i(\mathbf{x},\varepsilon) \geq 0$ for all $i \in \mathcal{N}$. If $\varepsilon \geq \sum_{h=1}^n a_h(\mu,\varepsilon)$ the non-negativity of the IEDV is always satisfied, when $\varepsilon < \sum_{h=1}^n a_h(\mu,\varepsilon)$ it is not anymore the case that $d_i(\mathbf{x},\varepsilon) \geq 0$ for all $a_i(\mu,\varepsilon) \geq 0$. From (30) it is clear that as the income share $\frac{x_i}{n\mu}$ of an individual in position i increases, for a given function $a_i(\mu,\varepsilon) \geq 0$ the IEDV may become negative. The higher possible income share consistent with the individual being in position i represents the upper bound of the condition requiring that $d_i(\mathbf{x},\varepsilon) \geq 0$ for all $i \in \mathcal{N}$. For an individual in position i the higher income share is obtained where all the $i-1$ poorer individuals have income 0, and the richer $n-i$ individuals have the same income x_i as individual i. It follows that for this distribution $x_i = \frac{n\mu}{n-i+1}$, for individual i the condition $d_i(\mathbf{x},\varepsilon) \geq 0$ implied by (30) becomes $a_i(\mu,\varepsilon) + \left[\varepsilon - \sum_{h=1}^n a_h(\mu,\varepsilon)\right] / (n - i + 1) \geq 0$ that is $\varepsilon + a_i(\mu,\varepsilon) \cdot (n - i + 1) \geq \sum_{h=1}^n a_h(\mu,\varepsilon)$. Since the term on the r.h.s. is independent from i then the condition is satisfied for all $i \in \mathcal{N}$ if it is satisfied for the individual i^* whose function $a_{i^*}(\mu,\varepsilon) \cdot (n - i^* + 1)$ gives the minimum value for a given pair $\varepsilon, \mu > 0$, that is the condition becomes for given $\varepsilon, \mu > 0$:

$$\varepsilon + \min{}_i \left\{a_i(\mu,\varepsilon) \cdot (n - i + 1) : i \in \mathcal{N}\right\} \geq \sum\nolimits_{h=1}^n a_h(\mu,\varepsilon).$$

Note that if $a_i(\mu,\varepsilon) = a(\mu,\varepsilon)$ for all $i \in \mathcal{N}$, then $\min_i \left\{a(\mu,\varepsilon) \cdot (n - i + 1) : i \in \mathcal{N}\right\}$ is obtained for $i = n$, that is the relevant distribution to consider is $(0, 0, ..., n\mu)$. ■

Proof of Proposition 1 (continued):

Note that according to (30) HE is not in general satisfied if $a_i(\mu,\varepsilon) \neq 0$ for some $i \in \mathcal{N}$. Thus if $\mathbf{d}(\mathbf{x},\varepsilon)$ satisfies condition LC, then $a_i(\mu,\varepsilon) = a(\mu,\varepsilon)$ $\forall i \in \mathcal{N}$ is a necessary and sufficient condition for satisfying HE. We obtain: $d_i(\mathbf{x},\varepsilon) = f_i(x_i, \mu,\varepsilon) = a(\mu,\varepsilon) + \frac{\varepsilon - na(\mu,\varepsilon)}{n\mu} x_i$, that is

$$d_i(\mathbf{x},\varepsilon) = a(\mu,\varepsilon)\left(1 - \frac{x_i}{\mu_x}\right) + \frac{x_i}{\mu_x}\frac{\varepsilon}{n}. \quad (35)$$

Note that since $d_i(\mathbf{x},\varepsilon)$ is differentiable w.r.t. ε, then also $a(\mu,\varepsilon)$ is differentiable w.r.t. ε. Note also that $d_i(\mathbf{x},\varepsilon) \geq \mathbf{0}$ implies that $a(\mu,\varepsilon) \geq 0$, moreover since we required $\mathbf{d}(\mathbf{x},\varepsilon) \geq 0 \; \forall \mathbf{x} \in \mathcal{X}^n$, then from (30) it must be that $a_i(\mu,\varepsilon) + \left[\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon)\right] \frac{x_i}{n\mu} \geq 0$, this condition has to be satisfied even for the extreme distribution $(0, 0, ...n\mu)$ implying that $\varepsilon \geq (n-1)a(\mu,\varepsilon)$ for $x_n = n\mu$, that is $(\varepsilon/n)\left[1 + 1/(n-1)\right] \geq a(\mu,\varepsilon)$. Furthermore the restriction in (32) requires that $\varepsilon/n + \mu \geq a(\mu,\varepsilon)$. We can therefore summarize that

Lemma 2 *An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC and HE if and only if there exists a set of continuous functions $a(\mu,\varepsilon) : \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}_{+}$ satisfying*

(I) $(\varepsilon/n)\left[1 + 1/(n-1)\right] \geq a(\mu,\varepsilon) \geq 0$ for all $\mu,\varepsilon > 0$, and

(II) $\varepsilon/n + \mu \geq a(\mu,\varepsilon) \geq 0$ for all $\mu,\varepsilon > 0$, such that for all $\mathbf{x} \in X^n$ and all $\varepsilon > 0$ the elements $d_i(\mathbf{x},\varepsilon)$ are given in (35) for all $i \in \mathcal{N}$.

Considering now condition VE, the only additional restriction imposed on (30) is that $b(\mu,\varepsilon) = \left[\varepsilon - \sum_{j=1}^{n} a_j(\mu,\varepsilon)\right]\frac{1}{n\mu} \geq 0$, i.e. $\varepsilon/n \geq a(\mu,\varepsilon)$ for all $\mu,\varepsilon > 0$. Since conditions HE and VE require that $\varepsilon/n \geq a(\mu,\varepsilon) \geq 0$, we can rewrite $a(\mu,\varepsilon)$ as a share of ε/n, that is $a(\mu,\varepsilon) = \left[1 - \gamma(\mu,\varepsilon)\right]\varepsilon/n$, where $1 \geq \gamma(\mu,\varepsilon) \geq 0$ for all $\mu,\varepsilon > 0$ i.e. $\gamma(\mu,\varepsilon) := \frac{\frac{\varepsilon}{n} - a(\mu,\varepsilon)}{\frac{\varepsilon}{n}}$. Given continuity of $a(\mu,\varepsilon)$ the function $\gamma(\mu,\varepsilon)$ is also continuous. Rearranging (35) we obtain:

$$d_i(\mathbf{x},\varepsilon) = f_i(x_i, \mu,\varepsilon) = \varepsilon \left\{ \gamma(\mu,\varepsilon)\frac{x_i}{n\mu_x} + \left[1 - \gamma(\mu,\varepsilon)\right]\frac{1}{n} \right\} \quad \forall i \in \mathcal{N}.$$

■

5.1.2 Proposition 2

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfies PI if and only if for $\gamma(\mu,\varepsilon)$ in (2) a continuous (or piecewise continuous) function $\eta : \mathbb{R}_{++} \to [0,1]$ exists and such that $\eta(\mu) := \lim_{\varepsilon \to 0^+} \gamma(\mu,\varepsilon)$ and such that for all $\mathbf{x} \in X^n$, and $\varepsilon > 0$

$$\mathbf{d}(\mathbf{x},\varepsilon) = (\mathbf{x} - \mu_x \mathbf{1}) \left\{ \exp\left[\int_{\mu_x}^{\mu_x + \frac{\varepsilon}{n}} \frac{\eta(\mu)}{\mu} d\mu \right] - 1 \right\} + \frac{\varepsilon}{n}\mathbf{1}. \tag{36}$$

In order to prove Proposition 2 we first prove that $\lim_{\varepsilon \to 0^+} \gamma(\mu,\varepsilon)$ exists for all IEDVs $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfying PI.

Lemma 3 *If an IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfies PI, then $\lim_{\varepsilon \to 0^+} \gamma(\mu,\varepsilon)$ exists and can be represented by a continuous (or piecewise continuous) function $\eta(\mu)$ where $\eta : \mathbb{R}_{++} \to [0,1]$.*

Proof of Lemma 3: We first prove that if $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ then the r.h.s. partial derivative of $\gamma(\mu,\varepsilon)$ w.r.t. ε is defined, finite and (piecewise) continuous in ε for all

$\varepsilon, \mu > 0$, then we will use this result to prove the lemma. Consider (2), denoting with $\boldsymbol{\alpha}$ the vector whose elements are $\alpha_i = \frac{x_i - \mu_x}{\mu_x}$, and rearranging, we obtain,

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\gamma\left(\mu,\varepsilon\right)\boldsymbol{\alpha} + \mathbf{1}\right]. \tag{37}$$

By property R we have that $\partial(\mathbf{x},\varepsilon^{+})$ is defined, finite and (piecewise) continuous in ε, from (37) we obtain

$$\partial(\mathbf{x},\varepsilon^{+}) = \frac{1}{n}\left[\gamma\left(\mu,\varepsilon\right)\boldsymbol{\alpha} + \mathbf{1}\right] + \frac{\partial\gamma\left(\mu,\varepsilon\right)}{\partial\varepsilon^{+}}\frac{\varepsilon}{n}\boldsymbol{\alpha}$$

where $\frac{\partial\gamma(\mu,\varepsilon)}{\partial\varepsilon^{+}}$ denotes the r.h.s. partial derivative of $\gamma(\mu,\varepsilon)$ w.r.t. ε. Since $\gamma\left(\mu,\varepsilon\right)$ is continuous in $\mu, \varepsilon > 0$, given the properties of $\partial(\mathbf{x},\varepsilon^{+})$ then also $\frac{\partial\gamma(\mu,\varepsilon)}{\partial\varepsilon^{+}}$ is defined, finite and (piecewise) continuous in ε for all $\varepsilon, \mu > 0$. Rewriting PI making use of this representation of the distributive rule in (37), we have:

$$\frac{\varepsilon}{n}\left[\gamma\left(\mu,\varepsilon\right)\boldsymbol{\alpha} + \mathbf{1}\right] + \frac{\varepsilon'}{n}\left[\gamma\left(\mu + \frac{\varepsilon}{n}, \varepsilon'\right)\boldsymbol{\beta} + \mathbf{1}\right] = \frac{\varepsilon + \varepsilon'}{n}\left[\gamma\left(\mu, \varepsilon + \varepsilon'\right)\boldsymbol{\alpha} + \mathbf{1}\right] \tag{38}$$

where, $\mathbf{y} = \mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon)$ therefore $\mu_y = \mu_x + \frac{\varepsilon}{n}$, and $\boldsymbol{\beta}$ is the vector of elements $\beta_i = \frac{y_i - \mu_y}{\mu_y}$. Then rewriting $\boldsymbol{\beta}$ in terms of $\boldsymbol{\alpha}$,

$$\begin{aligned}\boldsymbol{\beta} &= \frac{\mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon) - \mu\mathbf{1} - \frac{\varepsilon}{n}\mathbf{1}}{\mu + \frac{\varepsilon}{n}} = \frac{\mathbf{x} + \frac{\varepsilon}{n}\left[\gamma\left(\mu,\varepsilon\right)\boldsymbol{\alpha} + \mathbf{1}\right] - \mu\mathbf{1} - \frac{\varepsilon}{n}\mathbf{1}}{\mu + \frac{\varepsilon}{n}} \\ &= \frac{\mathbf{x} - \mu\mathbf{1} + \frac{\varepsilon}{n}\gamma\left(\mu,\varepsilon\right)\boldsymbol{\alpha}}{\mu + \frac{\varepsilon}{n}} = \boldsymbol{\alpha}\frac{\mu + \frac{\varepsilon}{n}\gamma\left(\mu,\varepsilon\right)}{\mu + \frac{\varepsilon}{n}}.\end{aligned} \tag{39}$$

Substituting into (38) and simplifying we get

$$\varepsilon\gamma\left(\mu,\varepsilon\right)\boldsymbol{\alpha} + \varepsilon'\gamma\left(\mu + \frac{\varepsilon}{n}, \varepsilon'\right)\boldsymbol{\alpha}\frac{\mu + \frac{\varepsilon}{n}\gamma\left(\mu,\varepsilon\right)}{\mu + \frac{\varepsilon}{n}} = \left(\varepsilon + \varepsilon'\right)\gamma\left(\mu, \varepsilon + \varepsilon'\right)\boldsymbol{\alpha}$$

that is for a generic $\alpha_i \neq 0$,

$$\varepsilon\gamma\left(\mu,\varepsilon\right) + \varepsilon'\gamma\left(\mu + \frac{\varepsilon}{n}, \varepsilon'\right)k = \left(\varepsilon + \varepsilon'\right)\gamma\left(\mu, \varepsilon + \varepsilon'\right) \tag{40}$$

where $k = \frac{\mu + \frac{\varepsilon}{n}\gamma(\mu,\varepsilon)}{\mu + \frac{\varepsilon}{n}} > 0$ since $\mu > 0, \varepsilon > 0$. Rearranging we obtain:

$$\gamma\left(\mu + \frac{\varepsilon}{n}, \varepsilon'\right)\frac{k}{\varepsilon} = \frac{\gamma\left(\mu, \varepsilon + \varepsilon'\right)}{\varepsilon} + \frac{\left[\gamma\left(\mu, \varepsilon + \varepsilon'\right) - \gamma\left(\mu,\varepsilon\right)\right]}{\varepsilon'}. \tag{41}$$

Consider now the limit for $\varepsilon' > 0$ approaching 0 of both sides of (41) that is:

$$\frac{k}{\varepsilon}\lim_{\varepsilon' \to 0^{+}}\gamma(\tilde{\mu}, \varepsilon') = \frac{\gamma\left(\mu,\varepsilon\right)}{\varepsilon} + \left.\frac{\partial\gamma\left(\mu,\varkappa\right)}{\partial\varkappa^{+}}\right|_{\varkappa = \varepsilon}$$

where $\tilde{\mu} = \mu + \frac{\varepsilon}{n} > 0$. Then, since $\varepsilon > 0$, $\frac{\gamma(\mu,\varepsilon)}{\varepsilon}$ is defined and finite, and $\frac{\partial\gamma(\mu,\varkappa)}{\partial\varkappa^+}$ has been shown to be defined and finite for all $\varepsilon, \mu > 0$, therefore, since $k > 0$, then $\lim_{\varepsilon'\to 0^+}\gamma(\tilde{\mu},\varepsilon')$ must be defined for all $\varepsilon, \tilde{\mu} > 0$. We define $\eta(\mu) := \lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon)$. It follows that $\eta(\tilde{\mu}) = \frac{\gamma(\mu,\varepsilon)}{k} + \frac{\varepsilon}{k}\cdot\left.\frac{\partial\gamma(\mu,\varkappa)}{\partial\varkappa^+}\right|_{\varkappa=\varepsilon}$, that is

$$\eta(\tilde{\mu}) = A(\mu,\varepsilon) + B(\mu,\varepsilon)\cdot\left.\frac{\partial\gamma(\mu,\varkappa)}{\partial\varkappa^+}\right|_{\varkappa=\varepsilon}$$

where $A(\mu,\varepsilon) = \frac{\gamma(\mu,\varepsilon)\mu+\gamma(\mu,\varepsilon)\frac{\varepsilon}{n}}{\mu+\frac{\varepsilon}{n}\gamma(\mu,\varepsilon)}$ and $B(\mu,\varepsilon) = \frac{\varepsilon\left[\mu+\frac{\varepsilon}{n}\right]}{\mu+\frac{\varepsilon}{n}\gamma(\mu,\varepsilon)}$. Given continuity of $\gamma(\mu,\varepsilon) \in [0,1]$, then also $A(\mu,\varepsilon)$ and $B(\mu,\varepsilon)$ are continuous in $\varepsilon, \mu > 0$.[38] Since $\left.\frac{\partial\gamma(\mu,\varkappa)}{\partial\varkappa^+}\right|_{\varkappa=\varepsilon}$ is (piecewise) continuous in ε for all $\varepsilon, \mu > 0$ and $\tilde{\mu} = \mu + \varepsilon/n$ then also $\eta(\mu)$ is (piecewise) continuous in $\tilde{\mu}$ for all $\tilde{\mu} > 0$.

Furthermore, by the definition of $\eta(\mu)$, given that $\gamma(\mu,\varepsilon)$ is bounded also $\eta(\mu)$ is bounded between $[0,1]$. ■

We can now prove Proposition 2.

Proof of Proposition 2:

Necessity: From Lemma 3 we know that PI is sufficient to ensure the existence of $\eta(\mu) := \lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon)$.

PI requires that, given two inequality equivalent distributions $\mathbf{x}, \mathbf{y} \in X^n$, there exists a unique sequence of distributions, whose average income is between μ_x and μ_y, that are inequality equivalent to these two distributions. Furthermore, the set of these distributions is independent from the starting distribution. Given that $\mathbf{d}(\mathbf{x},\varepsilon)$ is assumed to be continuous, then the sequence of inequality equivalent distributions should form a continuous path. We will try first to characterize this path considering the case in which PI is satisfied for "infinitesimal" and positive $\varepsilon, \varepsilon'$, then we will check that the solution applies for all $\varepsilon, \varepsilon' > 0$.

According to Proposition 1 we have that for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$

$$d_i(\mathbf{x},\varepsilon) = \varepsilon\left\{[1-\gamma(\mu,\varepsilon)]\frac{1}{n} + \gamma(\mu,\varepsilon)\frac{x_i}{n\mu_x}\right\} \quad \forall i \in \mathcal{N}, \tag{42}$$

which is the change in x_i associated with the sharing of surplus $\varepsilon > 0$. Thus, as $\varepsilon \to 0^+$, we obtain the infinitesimal increment in $\mathbf{x}$ which leaves inequality unchanged. Denote with $e = \varepsilon/n$ the equi-division of the surplus, let $\mathbf{d_n}(\mathbf{x},e) := \mathbf{d}(\mathbf{x},\varepsilon)$, and $\gamma_n(\mu,e) := \gamma(\mu,\varepsilon)$ for all $\mathbf{x} \in X^n$ all $\mu,\varepsilon > 0$, then $\frac{\mathbf{d_n}(\mathbf{x},e)}{e} = [1-\gamma_n(\mu,e)]\,\mathbf{1} + \gamma_n(\mu,e)\frac{\mathbf{x}}{\mu_x}$

[38] Notice that continuity of $\gamma(\mu,\varepsilon) \in [0,1]$ for $\varepsilon, \mu > 0$ is not sufficient to guarantee continuity of $\eta(\mu) := \lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon)$ for all $\mu > 0$. Here is an example:

$$\gamma(\mu,\varepsilon) = \begin{cases} 0 & \text{if} \quad 0 < \mu < \bar{\mu} \\ (\mu-\bar{\mu})/\varepsilon & \text{if} \quad \bar{\mu} \le \mu < \bar{\mu}+\varepsilon \\ 1 & \text{if} \quad \mu \ge \bar{\mu}+\varepsilon \end{cases}.$$

Note that $\gamma(\mu,\varepsilon)$ is continuous but $\eta(\mu) := \lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon) = \begin{cases} 0 & \text{if } 0 < \mu < \bar{\mu} \\ 1 & \text{if} \quad \mu \ge \bar{\mu} \end{cases}$ is discontinuous at $\mu = \bar{\mu} > 0$.

for all $\mathbf{x} \in X^n$. Notice that $\mu_y = \mu_x + e$, where the distribution $\mathbf{y}$ is obtained from $\mathbf{x}$ by sharing the surplus ε according to (42) , that is $d\mu = e$.

Considering

$$n \cdot \lim_{\varepsilon \to 0^+} \frac{\mathbf{d}(\mathbf{x}, \varepsilon)}{\varepsilon} = \lim_{d\mu \to 0^+} \frac{\mathbf{d_n}(\mathbf{x}, d\mu)}{d\mu} = \dot{\mathbf{x}}; \text{ where } \lim_{\varepsilon \to 0^+} \gamma(\mu,\varepsilon) = \lim_{e \to 0^+} \gamma_n(\mu,e) = \eta(\mu) \quad (43)$$

we can rearrange all possible equivalence relations as a system of dynamic equations. The system specifies the set of "infinitesimal" inequality equivalent changes in the distributions as the average income (or equivalently the amount of surplus to divide) increases. Thus, rearranging (42) we obtain

$$\dot{\mathbf{x}} = \eta(\mu_x) \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x} + \mathbf{1} \quad (44)$$

or equivalently

$$\dot{x}_i = \eta(\mu_x) \frac{x_i - \mu_x}{\mu_x} + 1 \quad \forall i \in \mathcal{N}. \quad (45)$$

Note that if we consider two distributions $\mathbf{x}_0 \in \mathcal{X}^n(\mu_0)$, $\mathbf{x}_1 \in \mathcal{X}^n(\mu_1)$, where $\mu_1 > \mu_0$, then according to this new representation $\mathbf{d}(\mathbf{x}_0, \varepsilon) = \int_{\mu_0}^{\mu_0+e} \dot{\mathbf{x}} d\mu$, and $\mathbf{d}(\mathbf{x}_1, \varepsilon') = \int_{\mu 1}^{\mu_1+e'} \dot{\mathbf{x}} d\mu$, therefore, if $\mu_0 + e = \mu_1$ then $\mathbf{d}(\mathbf{x}_0, \varepsilon + \varepsilon') = \int_{\mu_0}^{\mu_0+e+e'} \dot{\mathbf{x}} d\mu = \int_{\mu_0}^{\mu_0+e} \dot{\mathbf{x}} d\mu + \int_{\mu 1}^{\mu_1+e'} \dot{\mathbf{x}} d\mu = \mathbf{d}(\mathbf{x}_0, \varepsilon) + \mathbf{d}(\mathbf{x}_1, \varepsilon')$, which is exactly the Path Independence condition. Note that the integral symbol applied denotes the sum of all integrals defined over all disjoint sub-intervals of $[\mu_x, \mu_x + e]$ where $\eta(\mu)$ is continuous.

In order to make explicit the equivalence criterion, implicitly defined in (44) , we need to solve the system of differential, non-autonomous, equations in (45) in terms of μ. All the equations are independent, thus they can be solved independently. Denote $z_i = x_i - \mu_x$; noticing that $\dot{z}_i = \dot{x}_i - 1$ we can rewrite (45) as

$$\dot{z}_i(\mu) = \eta(\mu) \frac{z_i(\mu)}{\mu}. \quad (46)$$

Define $\frac{\eta(\mu)}{\mu} = g(\mu) \to \dot{z}_i = g(\mu) z_i$, multiply both sides of this equation by $\exp(-\int_{\mu_0}^{\mu} g(\tau) d\tau)$, and let $\exp(-\int_{\mu_0}^{\mu} g(\tau) d\tau) \cdot z_i(\mu) = h_i(\mu)$. It follows that :

$$\dot{h}_i(\mu) = -g(\mu) \exp\left(-\int_{\mu_0}^{\mu} g(\tau) d\tau\right) \cdot z_i(\mu) + \exp\left(-\int_{\mu_0}^{\mu} g(\tau) d\tau\right) \cdot \dot{z}_i(\mu). \quad (47)$$

Since $\dot{z}_i = g(\mu) z_i$, $\dot{h}_i(\mu) = 0$, thus

$$\exp\left(-\int_{\mu_0}^{\mu} g(\tau) d\tau\right) \cdot z_i(\mu) = k. \quad (48)$$

That is, in terms of the distances of x_i from the mean income:

$$x_i - \mu_x = k \exp\left(\int_{\mu_0}^{\mu_x} g(\tau) d\tau\right) \quad \text{for all } i \in \mathcal{N}. \quad (49)$$

Comparing with distribution $\mathbf{y}$, for which $y_i - \mu_y = k \exp(\int_{\mu_0}^{\mu_y} g(\tau) d\tau)$ we obtain:

$$\frac{(y_i - \mu_y)}{(x_i - \mu_x)} = \exp\left(\int_{\mu_x}^{\mu_y} \frac{\eta(\mu)}{\mu} d\mu\right) \quad \text{for all } i \in \mathcal{N}. \tag{50}$$

Which could be rearranged as in Corollary 1

$$x_i - \mu_x = \left(y_i - \mu_y\right) \exp\left(-\int_{\mu_x}^{\mu_y} \frac{\eta(\mu)}{\mu} d\mu\right) \qquad \text{for all } i \in \mathcal{N}. \tag{51}$$

Which is the distributive rule $\mathbf{d}(\mathbf{x}, \varepsilon)$ associated with the equivalence criterion represented by (5)? Considering the relations $\mu_y = \mu_x + \frac{\varepsilon}{n}$ and $\mathbf{x} + \mathbf{d}(\mathbf{x}, \varepsilon) = \mathbf{y}$ and rearranging (2) , we can obtain

$$\mathbf{d}(\mathbf{x}, \varepsilon) = \frac{\varepsilon}{n} \left\{ \gamma(\mu,\varepsilon) \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x} + \mathbf{1} \right\} \tag{52}$$

that is $\mathbf{y} - \mathbf{x} = \mathbf{d}(\mathbf{x}, \varepsilon) = \left(\mu_y - \mu_x\right) \left\{ \gamma(\mu,\varepsilon) \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x} + \mathbf{1} \right\}$, which could be rearranged as

$$\left(\mathbf{y} - \mu_y \mathbf{1}\right) = (\mathbf{x} - \mu_x \mathbf{1}) \left\{ 1 + \gamma(\mu,\varepsilon) \frac{\mu_y - \mu_x}{\mu_x} \right\}. \tag{53}$$

Comparing this equation with (5) we obtain $\exp(\int_{\mu_x}^{\mu_y} \frac{\eta(\tau)}{\tau} d\tau)) = 1 + \gamma(\mu,\varepsilon) \frac{\mu_y - \mu_x}{\mu_x}$ which allows us to specify the equation for $\gamma(\mu,\varepsilon)$:

$$\left\{ \exp[\int_{\mu_x}^{\mu_x + \frac{\varepsilon}{n}} \frac{\eta(\tau)}{\tau} d\tau] - 1 \right\} \mu_x \frac{n}{\varepsilon} = \gamma(\mu,\varepsilon) \tag{54}$$

thus, equivalently the associated IEDVs are:

$$\mathbf{d}(\mathbf{x}, \varepsilon) = (\mathbf{x} - \mu_x \mathbf{1}) \left\{ \exp[\int_{\mu_x}^{\mu_x + \frac{\varepsilon}{n}} \frac{\eta(\tau)}{\tau} d\tau] - 1 \right\} + \frac{\varepsilon}{n} \mathbf{1}. \tag{55}$$

Sufficiency: Check that (36) satisfies conditions VE, HE, LC and PI. ■

5.1.3 Proposition 3

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{P}}(\mathbf{x}, \varepsilon)$ satisfies B if and only if either there exist constants $\lambda \in (0, 1]$ and $\tilde{\theta} \geq 0$ such that for all $\mathbf{x} \in \mathcal{X}^n$ either

$$\mathbf{d}(\mathbf{x}, \varepsilon) = (\mathbf{x} - \mu_x \mathbf{1}) \left\{ \left[\exp(\int_{\mu_x}^{\mu_x + \frac{\varepsilon}{n}} \frac{\lambda}{\mu + \tilde{\theta}} d\mu) \right] - 1 \right\} + \frac{\varepsilon}{n} \mathbf{1}. \tag{56}$$

or $\mathbf{d}(\mathbf{x}, \varepsilon) = \frac{\varepsilon}{n} \mathbf{1}$.

Proof of Proposition 3:

This proof is divided into 3 steps. We first prove that the set of all possible distributions such that B applies for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{P}}(\mathbf{x},\varepsilon)$ lies on the rays w.r.t. an appropriate origin $-\tilde{\theta}\mathbf{1}$. Then we show what are the restrictions on $\eta(\mu)$ implied by this property, and that $-\tilde{\theta}\mathbf{1} \leq \mathbf{0}$. Finally we apply this characterization of $\eta(\mu)$ to the result of Proposition 2 in order to characterize the equivalence criterion in Corollary *2*. Then, according to this result, we characterize the IEDV.

Step 1:

Define by $\mathcal{D}'_{\mathcal{E}}(\mathbf{x},\varepsilon) := \{\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ such that $\lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon) = \eta(\mu)$ exists$\}$.

Lemma 4 *Consider a distributive rule* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}'_{\mathcal{E}}(\mathbf{x},\varepsilon)$ *satisfying B. If* $\lim_{\varepsilon\to 0^+}\gamma(\mu,\varepsilon) = \eta(\mu) = 0$ *for some distribution* $\mathbf{x} \in \mathcal{X}^n$*, then* $\eta(\hat{\mu}) = \hat{\eta} = 0$ *for all* $\hat{\mathbf{x}} \in \mathcal{X}^n$.

Proof of Lemma 4: According to (43) and (44) it follows that

$$\boldsymbol{\delta}(\mathbf{x}) = \frac{1}{n}\left[\eta(\mu_x)\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x} + \mathbf{1}\right], \tag{57}$$

where, given Lemma 3 $\eta(\mu_x) \in [0,1]$ exists. If $\eta(\mu_x) = 0$, then $\boldsymbol{\delta}(\mathbf{x}) = \frac{1}{n}\mathbf{1}$. Notice that all distributions $\mathbf{y} = \mu_y\mathbf{1}$ satisfy the condition $\boldsymbol{\delta}(\mathbf{y}) = \frac{1}{n}\mathbf{1}$. Then, according to axiom B, it must be true that $\boldsymbol{\delta}(\tilde{\mathbf{x}}) = \frac{1}{n}\mathbf{1}$ for all distributions $\tilde{\mathbf{x}} = \beta\mathbf{x} + (1-\beta)\mathbf{y}$ $\forall\beta \in [0,1]$. Therefore, since $\mathbf{x}$ could be different from $\mu_x\mathbf{1}$, then, if $\beta \in (0,1]$, also $\tilde{\mathbf{x}} \neq \tilde{\mu}\mathbf{1}$ where $\tilde{\mu} = \beta\mu_x + (1-\beta)\mu_y$. Thus in order to satisfy $\boldsymbol{\delta}(\tilde{\mathbf{x}}) = \frac{1}{n}\mathbf{1}$ it must be that $\eta(\tilde{\mu}_x) = 0$; in other words $\eta(\mu) = 0$ for all $\mu > 0$. ■

For $\mu, \bar{\mu}, \hat{\mu} > 0$, let $\eta := \eta(\mu)$, $\hat{\eta} := \eta(\hat{\mu})$ and $\bar{\eta} := \eta(\bar{\mu})$.

Lemma 5 *Suppose* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}'_{\mathcal{E}}(\mathbf{x},\varepsilon)$ *for all* $\mathbf{x} \in X^n$*. Consider two distributions* $\mathbf{x} \in \mathcal{X}^n(\mu)$ *and* $\hat{\mathbf{x}} \in \mathcal{X}^n(\hat{\mu})$*, where* $\mu \leq \hat{\mu}$*, and* $\mathbf{d}(\mathbf{x},\varepsilon)$ *satisfies B, such that* $\boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$*. Then, either (a) the distribution* $\bar{\mathbf{x}} \in \mathcal{X}^n(\bar{\mu})$ *such that* $\delta(\mathbf{x}) = \delta(\bar{\mathbf{x}})$ *and either* $\bar{\mu} \leq \mu$ *or* $\hat{\mu} \leq \bar{\mu}$ *lies on the same ray in which* $\mathbf{x}$ *and* $\hat{\mathbf{x}}$ *lie, or (b)* $\eta = \hat{\eta} = \bar{\eta} = 0$.

Proof of Lemma 5: Consider (57), then $\boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$ is equivalent to

$$\eta\frac{\mathbf{x}-\mu\mathbf{1}}{\mu} = \hat{\eta}\frac{\hat{\mathbf{x}}-\hat{\mu}\mathbf{1}}{\hat{\mu}}. \tag{58}$$

Suppose that there exists an $\bar{\mathbf{x}} \in \mathcal{X}^n(\bar{\mu})$ such that $\eta\frac{\mathbf{x}-\mu\mathbf{1}}{\mu} = \bar{\eta}\frac{\bar{\mathbf{x}}-\bar{\mu}\mathbf{1}}{\bar{\mu}}$ and w.l.o.g. suppose $\bar{\mu} \leq \mu$, then by axiom B, for all distributions $\tilde{\mathbf{x}} \in \mathcal{X}^n$ such that $\tilde{\mathbf{x}} = \beta\bar{\mathbf{x}} + (1-\beta)\hat{\mathbf{x}}$, $\beta \in [0,1]$ it must be $\boldsymbol{\delta}(\tilde{\mathbf{x}}) = \boldsymbol{\delta}(\mathbf{x})$. Thus if $\tilde{\mathbf{x}}$ and $\mathbf{x}$ do not lie on the same ray it is the case that there exists a distribution $\tilde{\mathbf{x}}$, such that $\tilde{\mu} = \mu$ and $\tilde{\mathbf{x}} \neq \mathbf{x}$. Then since $\boldsymbol{\delta}(\tilde{\mathbf{x}}) = \boldsymbol{\delta}(\mathbf{x})$ it must be that $\eta = \tilde{\eta} = 0$. Thus, from Lemma 4 we know that $\eta = 0$ for every $\mathbf{x} \in X^n$. ■

Lemma 6 *Suppose* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}'_{\mathcal{E}}(\mathbf{x},\varepsilon)$ *for all* $\mathbf{x} \in X^n$*. Consider two distributions* $\mathbf{x} \in \mathcal{X}^n(\mu)$ *and* $\hat{\mathbf{x}} \in \mathcal{X}^n(\hat{\mu})$*, where* $\mu < \hat{\mu}$ *and* $\boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$*. If* $\eta \neq 0$ *all possible distributions such that B applies lie on the rays w.r.t. the origin* $-\tilde{\theta}\mathbf{1}$*, for*

$$\tilde{\theta} = \frac{\hat{\eta}-\eta}{\frac{\eta}{\mu}-\frac{\hat{\eta}}{\hat{\mu}}}. \tag{59}$$

Proof of Lemma 6: From Lemmata 4 and 5 we know that if $\eta \neq 0$, also $\hat{\eta} \neq 0$, and that all distributions $\tilde{\mathbf{x}}$ such that $\boldsymbol{\delta}(\tilde{\mathbf{x}}) = \boldsymbol{\delta}(\mathbf{x}) = \boldsymbol{\delta}(\hat{\mathbf{x}})$ lie on the ray identified by $\mathbf{x}, \hat{\mathbf{x}}$. Given $\mathbf{x}, \hat{\mathbf{x}}$, there always exists an origin $-\tilde{\theta}\mathbf{1}$, such that these two distributions lie on a ray through this point. We need to show which is the value of $\tilde{\theta}$. The two distributions must satisfy the following relations:

$$\begin{cases} \eta \frac{\mathbf{x}-\mu\mathbf{1}}{\mu} = \hat{\eta} \frac{\hat{\mathbf{x}}-\hat{\mu}\mathbf{1}}{\hat{\mu}} \\ \frac{\mathbf{x}-\mu\mathbf{1}}{\mu+\tilde{\theta}} = \frac{\hat{\mathbf{x}}-\hat{\mu}\mathbf{1}}{\hat{\mu}+\tilde{\theta}} \end{cases} \tag{60}$$

that is, substituting,

$$\frac{\eta}{\mu}(\mu + \tilde{\theta}) = \frac{\hat{\eta}}{\hat{\mu}}(\hat{\mu} + \tilde{\theta}). \tag{61}$$

Rearranging we obtain $\tilde{\theta} = \frac{\hat{\eta}-\eta}{\frac{\eta}{\mu}-\frac{\hat{\eta}}{\hat{\mu}}}$. Since the value of $\tilde{\theta}$ does not depend explicitly on the two distributions, but only on their means and sharing parameters, this means that once we identify two distributions satisfying B, this is enough to characterize all the relationships between all the other distributions in X^n. ■

Proof of Proposition 3 (continued)

Step 2:

Lemma 7 *For all* $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}'_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ *satisfying B, if* $\eta(\mu) \neq 0$ *then* $\eta(\mu) = \lambda \frac{\mu}{\mu+\tilde{\theta}}$ *where* $\tilde{\theta} \geq 0$ *is defined by* (59), *and* $1 \geq \lambda > 0$.

Proof of Lemma 7: From (61) we know that once we are able to define a value for $\tilde{\theta}$, then it must be the case that at different values of average income the function $\eta(\mu)$ satisfies the following functional relation

$$\frac{\eta(\mu)}{\mu}(\mu + \tilde{\theta}) = \frac{\eta(\hat{\mu})}{\hat{\mu}}(\hat{\mu} + \tilde{\theta}) \quad \forall \mu, \hat{\mu} > 0.$$

Let $\phi(\mu) = \frac{\eta(\mu)}{\mu}\left(\mu + \tilde{\theta}\right) \neq 0$. Then it must be that:

$$\phi(\mu) = \phi(\hat{\mu}) \quad \forall \mu, \hat{\mu} > 0.$$

Thus $\phi(\mu)$ must be constant, that is, there exists a real number $\lambda \neq 0$ such that $\phi(\mu) = \lambda$ for all $\mu > 0$. Substituting and rearranging we obtain:

$$\eta(\mu) = \lambda \frac{\mu}{(\mu + \tilde{\theta})}\ .$$

Since $\eta(\mu)$ is supposed to be bounded between $(0, 1]$ for every $\mu > 0$, then it must be that $\theta \geq 0$, and $\lambda \in (0, 1]$. Suppose $\tilde{\theta} < 0$, then in order to have $\eta(\mu) > 0$ it has to be $\lambda > 0$ if $\mu + \tilde{\theta} > 0$, and $\lambda < 0$ if $\mu + \tilde{\theta} < 0$, that contradicts the fact that λ is independent from μ. Consider $\tilde{\theta} \geq 0$, then since we supposed $\eta(\mu) > 0$ we get $\lambda > 0$, and in order to satisfy $\eta(\mu) \leq 1$ it should be $\lambda \leq \frac{\mu+\tilde{\theta}}{\mu}$, $\forall \mu > 0, \tilde{\theta} \geq 0$; which is satisfied only if $\lambda \leq 1$. ■

Proof of Proposition 3 (continued).

Once we consider the case of $\eta(\mu) = 0$ for some $\mu > 0$ then, from Lemma 4 for all $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}'_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ satisfying B we have that $\eta(\mu) = 0$ $\forall \mu > 0$. This condition together with the result of Lemma 6 allows to define:

$$\eta(\mu) = \lambda \frac{\mu}{(\mu + \tilde{\theta})} \quad \forall \mu > 0, \tilde{\theta} \geq 0, \lambda \in [0, 1] . \tag{62}$$

Step 3:

This last part of the proof consists in simply realizing that from Lemma 3, adding PI restricts the class of admissible $\mathbf{d}(\mathbf{x}, \varepsilon)$ from $\mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ to $\mathcal{D}'_{\mathcal{E}}(\mathbf{x}, \varepsilon)$. Adding B we characterize $\eta(\mu) = \lambda \frac{\mu}{(\mu + \tilde{\theta})}$, then substituting it into (5) we obtain

$$x_i - \mu_x = (y_i - \mu_y) \exp(-\int_{\mu_x}^{\mu_y} \frac{\lambda}{(\mu + \tilde{\theta})} d\mu), \tag{63}$$

that could be solved noticing that $\int_{\mu_x}^{\mu_y} \frac{-\lambda}{(\mu + \tilde{\theta})} d\mu = \lambda \left[\ln(\mu_x + \tilde{\theta}) - \ln(\mu_y + \tilde{\theta}) \right]$ leading to:

$$\frac{y_i - \mu_y}{(\mu_y + \tilde{\theta})^\lambda} = \frac{x_i - \mu_x}{(\mu_x + \tilde{\theta})^\lambda}$$

as in Corollary 2. According to the same line of reasoning it is possible to characterize the IEDVs substituting from (62) in (36) . ■

5.1.4 Proposition 4

An IEDV $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ satisfies EB if and only if there exist two continuous functions $\omega : \mathbb{R}_{++} \to [0, 1]$, $\theta : \mathbb{R}_{++} \to \mathbb{R}_+$ such that for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$:

$$\mathbf{d}(\mathbf{x}, \varepsilon) = \frac{\varepsilon}{n} \left(\omega(\varepsilon) \cdot \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x + \theta(\varepsilon)} + \mathbf{1} \right) . \tag{64}$$

Proof of Proposition 4:

Sufficiency: Check that (64) satisfies conditions LC, HE, VE, EB.

Necessity: Consider the characterization of the IEDVs $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ in (2)

$$\mathbf{d}(\mathbf{x}, \varepsilon) = \frac{\varepsilon}{n} \left\{ \gamma(\mu_x, \varepsilon) \cdot \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x} + \mathbf{1} \right\} .$$

satisfying EB. Following the same line of reasoning as in Lemma 6, given two distributions such that $\mathbf{d}(\mathbf{x}, \varepsilon) = \mathbf{d}(\hat{\mathbf{x}}, \varepsilon)$ for some $\varepsilon > 0$, then also all distributions exhibiting the same distributive rule for a given $\varepsilon > 0$ must lie on the ray through the origin $\tilde{\theta}$ which may depend on ε. This means

$$\begin{cases} \gamma(\mu_x, \varepsilon) \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x} = \gamma(\mu_y, \varepsilon) \frac{\mathbf{x} - \mu_y \mathbf{1}}{\mu_y} \\ \frac{\mathbf{x} - \mu_x \mathbf{1}}{\mu_x + \theta(\varepsilon)} = \frac{\mathbf{x} - \mu_y \mathbf{1}}{\mu_y + \theta(\varepsilon)} \end{cases}$$

that is

$$\gamma(\mu_x,\varepsilon)\frac{\mu_x+\theta(\varepsilon)}{\mu_x} = \gamma(\mu_y,\varepsilon)\frac{\mu_y+\theta(\varepsilon)}{\mu_y}. \tag{65}$$

Letting $\varpi(\mu,\varepsilon) = \gamma(\mu,\varepsilon)\frac{\mu+\theta(\varepsilon)}{\mu}$, it becomes evident that

$$\varpi(\mu_x,\varepsilon) = \varpi(\mu_y,\varepsilon)\ \forall \mu_x,\mu_y > 0, \forall \varepsilon > 0.$$

This means that $\varpi(\mu,\varepsilon)$ does not depend on μ, therefore there exists a function $\omega : \mathbb{R}_{++} \to [0,1]$ such that $\varpi(\mu,\varepsilon) = \omega(\varepsilon)\ \forall \mu > 0, \varepsilon > 0$. Substituting into (65) we obtain:

$$\gamma(\mu,\varepsilon) = \omega(\varepsilon)\frac{\mu}{\mu+\theta(\varepsilon)}\ \forall \mu,\varepsilon > 0. \tag{66}$$

Given the restrictions imposed on $\gamma(\mu,\varepsilon)$, then $\omega(\varepsilon)$ and $\theta(\varepsilon)$ are continuous in ε, furthermore $\theta(\varepsilon) \geq 0$, and $1 \geq \omega(\varepsilon) \geq 0$. Then we obtain (64) . ■

5.1.5 Proposition 5

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfies PI and EB if and only if there exist $\tilde{\theta} \geq 0$ such that for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$

$$\text{either } \mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left(\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x+\tilde{\theta}}+\mathbf{1}\right) \tag{67}$$

$$\text{or } \mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}. \tag{68}$$

Proof of Proposition 5:

Sufficiency: Check that the obtained IEDVs satisfy conditions LC, HE, VE, PI, EB.

Necessity: We identify the restrictions imposed by PI on the characterization in Proposition 4.

We can express PI for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ as in (40) . Denoting with $e := \varepsilon/n$, $\gamma_n(\mu,e) := \gamma(\mu,\varepsilon)$, $\omega_n(e) := \omega(\varepsilon)$, and $\theta_n(e) := \theta(\varepsilon)$ for all $\varepsilon > 0$, where $\omega_n(e) \in [0,1]$ and $\theta_n(e) \geq 0$ for all $e > 0$, we have

$$e\gamma_n(\mu,e) + e'\gamma_n(\mu+e,e')\frac{\mu+e\gamma_n(\mu,e)}{\mu+e} = (e+e')\gamma_n(\mu,e+e').$$

Substituting for $\gamma_n(\mu,e) = \omega_n(e)\frac{\mu}{\mu+\theta_n(e)}$, and simplifying we obtain:

$$\frac{e\omega_n(e)}{\mu+\theta_n(e)} + \frac{e'\omega_n(e')}{\mu+e+\theta_n(e')}\left[1+e\frac{\omega_n(e)}{\mu+\theta_n(e)}\right] = \frac{(e+e')\,\omega_n(e+e')}{\mu+\theta_n(e+e')} \tag{69}$$

which is a functional equation for $\theta_n(e)$ and $\omega_n(e)$ that is satisfied for all $\mu, e, e' > 0$. Note that if we swap e and e' then the r.h.s. of (69) is not affected while the l.h.s. changes. Equating it with the l.h.s. of (69) gives a new condition implied by (69) :

$$\frac{e\omega_n(e)}{\mu+\theta_n(e)} + \frac{e'\omega_n(e')}{\mu+e+\theta_n(e')}\left[1+\frac{e\omega_n(e)}{\mu+\theta_n(e)}\right]$$
$$= \frac{e'\omega_n(e')}{\mu+\theta_n(e')} + \frac{e\omega_n(e)}{\mu+e'+\theta_n(e)}\left[1+\frac{e'\omega_n(e')}{\mu+\theta_n(e')}\right] \tag{70}$$

that is satisfied for all $\mu, e, e' > 0$. We solve now (70) for $\theta_n(e)$ and $\omega_n(e)$ and then check whether the solutions obtained satisfy (69). Rearranging (70) we get

$$\begin{aligned} & \left[\frac{e\omega_n(e)}{\mu+\theta_n(e)} - \frac{e\omega_n(e)}{\mu+e'+\theta_n(e)}\right] - \left[\frac{e'\omega_n(e')}{\mu+\theta_n(e')} - \frac{e'\omega_n(e')}{\mu+e+\theta_n(e')}\right] \\ = \; & [e'\omega_n(e') \cdot e\omega_n(e)] \left[\frac{1}{[\mu+e'+\theta_n(e)][\mu+\theta_n(e')]} - \frac{1}{[\mu+e+\theta_n(e')][\mu+\theta_n(e)]}\right], \end{aligned} \tag{71}$$

after solving within the brackets and dividing for $e', e > 0$ we obtain

$$\begin{aligned} & \frac{\omega_n(e)}{[\mu+\theta_n(e)][\mu+e'+\theta_n(e)]} - \frac{\omega_n(e')}{[\mu+\theta_n(e')][\mu+e+\theta_n(e')]} \\ = \; & [\omega_n(e') \cdot \omega_n(e)] \left[\frac{e\mu + e\theta_n(e) - e'\mu - e'\theta_n(e')}{[\mu+e'+\theta_n(e)][\mu+\theta_n(e')][\mu+e+\theta_n(e')][\mu+\theta_n(e)]}\right]. \end{aligned} \tag{72}$$

After multiplying both sides by the denominator of the r.h.s. we have

$$\begin{aligned} & \omega_n(e)[\mu+\theta_n(e')][\mu+e+\theta_n(e')] - \omega_n(e')[\mu+e'+\theta_n(e)][\mu+\theta_n(e)] \\ = \; & [\omega_n(e') \cdot \omega_n(e)][e\mu + e\theta_n(e) - e'\mu - e'\theta_n(e')] \end{aligned} \tag{73}$$

that can be rewritten as a second degree (functional) equation in terms of μ:

$$\mu^2 f(e,e') + \mu g(e,e') + k(e,e') = 0 \quad \text{for all } \mu, e, e' > 0$$

where

$$f(e,e') \; : \; = \omega_n(e) - \omega_n(e') \tag{74}$$

$$\begin{aligned} g(e,e') \; : \; & = \omega_n(e)[2\theta_n(e') + e] - \omega_n(e')[2\theta_n(e) + e'] \\ & - [\omega_n(e') \cdot \omega_n(e)][e - e'] \end{aligned} \tag{75}$$

$$\begin{aligned} k(e,e') \; : \; & = \omega_n(e)\theta_n(e')[e + \theta_n(e')] - \omega_n(e')\theta_n(e)[e' + \theta_n(e)] \\ & - [\omega_n(e') \cdot \omega_n(e)][e\theta_n(e) - e'\theta_n(e')]. \end{aligned} \tag{76}$$

Since $f(e,e')$, $g(e,e')$, and $k(e,e')$ do not depend on μ, the functional equation is satisfied if and only if $f(e,e') = g(e,e') = k(e,e') = 0$ for all $e, e' > 0$. The first two functions must be equal to 0 for all $e, e' > 0$ because otherwise it is impossible to eliminate μ from the quadratic equation, therefore also $k(e,e')$ must be equal to 0 for all $e, e' > 0$.

From $f(e,e') = 0$ for all $e, e' > 0$ we know that $\omega_n(e) = \omega_n(e')$ for all $e, e' > 0$. That is $\omega_n(e)$ must be constant, i.e. there exists a real number $\tilde{\omega} \in [0,1]$ such that

$$\omega_n(e) = \tilde{\omega} \in [0,1] \quad \forall e > 0. \tag{77}$$

Thus, after substituting from (77) into (75) we obtain for the condition $g(e,e') = 0$ for all $e, e' > 0$:

$$\begin{aligned} g(e,e') \; = \; & 0 \leftrightarrow \tilde{\omega}[2\theta_n(e') + e] - \tilde{\omega}[2\theta_n(e) + e'] = (\tilde{\omega})^2 [e - e'] \\ \leftrightarrow \; & \tilde{\omega} 2[\theta_n(e') - \theta_n(e)] + \tilde{\omega}[e - e'] = (\tilde{\omega})^2 [e - e'] \\ \leftrightarrow \; & \tilde{\omega} 2[\theta_n(e) - \theta_n(e')] = \tilde{\omega} \cdot (1 - \tilde{\omega})[e - e'] \qquad \text{for all } e, e' > 0. \end{aligned}$$

That is either $\tilde{\omega} = 0$ or, if $\tilde{\omega} \in (0,1]$, the functional equation $g(e,e') = 0$ for all $e, e' > 0$ can be simplified as

$$\theta_n(e) - \theta_n(e') = (1-\tilde{\omega})\,\frac{e-e'}{2} \qquad \text{for all } e, e' > 0,$$

that is

$$\theta_n(e) - (1-\tilde{\omega})\,\frac{e}{2} = \theta_n(e') - (1-\tilde{\omega})\,\frac{e'}{2} \qquad \text{for all } e, e' > 0.$$

It follows that $\theta_n(e) - (1-\tilde{\omega})\,\frac{e}{2}$ is constant for all $e > 0$, that is

$$\theta_n(e) = \theta + (1-\tilde{\omega})\,\frac{e}{2} \quad \text{for all } e > 0 \tag{78}$$

where $\theta \geq 0$, and $\tilde{\omega} \in (0,1]$. The restriction on the arbitrary constant θ is introduced in order to satisfy the initial condition $\theta_n(e) \geq 0$ for all $e > 0$. Given (77) the solution of (75) is:

$$\begin{aligned} \text{either } \theta_n(e) &= \theta + (1-\tilde{\omega})\,\frac{e}{2}, \text{ where } \theta \geq 0, \quad \text{for all } e > 0 \ \text{ if } \tilde{\omega} \in (0,1], \quad (79)\\ \text{or} \qquad \tilde{\omega} &= 0. \end{aligned}$$

Note that $\tilde{\omega} = 0$ is a solution also for the functional equation $k(e,e') = 0$ for all $e, e' > 0$. While if $\tilde{\omega} \in (0,1]$ (76) imposes additional restrictions on (78). After substituting from (77) into the condition requiring $k(e,e') = 0$ for all $e, e' > 0$ we obtain

$$[\theta_n(e)]^2 + e'\theta_n(e) - [\theta_n(e')]^2 - e\theta_n(e') = -\tilde{\omega}\,[e\theta_n(e) - e'\theta_n(e')] \quad \text{for all } e, e' > 0.$$

Adding on both sides $e'\theta_n(e') - e\theta_n(e)$ and simplifying we get

$$[\theta_n(e) - \theta_n(e')]\,[\theta_n(e) + \theta_n(e')] + [\theta_n(e) + \theta_n(e')]\,[e' - e] = (1+\tilde{\omega})\,[e'\theta_n(e') - e\theta_n(e)]$$

for all $e, e' > 0$, that is

$$[\theta_n(e) - \theta_n(e') - e + e']\,[\theta_n(e) + \theta_n(e')] = (1+\tilde{\omega})\,[e'\theta_n(e') - e\theta_n(e)] \quad \text{for all } e, e' > 0.$$

Substituting from (78) we obtain

$$\begin{aligned} &\tilde{\omega}\left[\left(\frac{1-\tilde{\omega}}{2} - 1\right)(e-e')\right] \cdot \left[2\theta + (1-\tilde{\omega})\,\frac{e+e'}{2}\right] \\ = \ &(1+\tilde{\omega})\left[e'\theta + (1-\tilde{\omega})\,\frac{(e')^2}{2} - e\theta - (1-\tilde{\omega})\,\frac{e^2}{2}\right] \quad \text{for all } e, e' > 0, \end{aligned}$$

that is

$$\begin{aligned} &-\tilde{\omega}\left[\left(\frac{1+\tilde{\omega}}{2}\right)(e-e')\right] \cdot \left[2\theta + (1-\tilde{\omega})\,\frac{e+e'}{2}\right] \\ = \ &(1+\tilde{\omega})\left[(e'-e)\,\theta + (1-\tilde{\omega})\,\frac{(e'-e)\,(e'+e)}{2}\right] \quad \text{for all } e, e' > 0. \end{aligned}$$

This functional equation can be simplified giving

$$\tilde{\omega}\left(1-\tilde{\omega}\right)\frac{e+e'}{4}=\left(1-\tilde{\omega}\right)\left[\theta+\frac{e'+e}{2}\right]\quad\text{for all } e,e'>0. \tag{80}$$

The condition $\tilde{\omega}=1$ is a solution for (80). If $\tilde{\omega}\in(0,1)$ we can simplify (80) as

$$\left(\frac{\tilde{\omega}}{4}-\frac{1}{2}\right)(e+e')=\theta\ \text{ for all } e,e'>0. \tag{81}$$

Since the r.h.s. is independent from $e,e'>0$ then also the l.h.s. has to be independent, this is only possible if $\tilde{\omega}=2$ but we have supposed that $\tilde{\omega}\in(0,1)$. As a result the general solution of (80) is $\tilde{\omega}=1$, and $\theta\geq 0$. Substituting $\tilde{\omega}=1$ into (78) we get $\theta_n(e)=\theta\geq 0$. Notice that the solutions (i) $\omega_n(e)=0$ for all $e>0$, and (ii) $\theta_n(e)=\theta\geq 0$ and $\omega_n(e)=1$ for all $e>0$ satisfy (69). We can therefore conclude that the general solution of (69) is:

$$\begin{array}{llcl}\text{either } (i) & \omega_n(e) & = & 0 \qquad \text{for all } e>0,\\ \text{or} \qquad (ii) & \theta_n(e) & = & \theta\geq 0 \ \text{ and } \ \omega_n(e)=1 \quad \text{for all } e>0\end{array}$$

Thus it must be that

$$\text{either } \gamma_n(\mu,e)=0 \quad \text{or} \quad \gamma_n(\mu,e)=\frac{\mu}{\mu+\theta}\quad \text{for all } \mu,e>0,$$

for all of all $\mathbf{d}(\mathbf{x},\varepsilon)\in\mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfying PI+EB. ■

5.1.6 Proposition 6

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon)\in\mathcal{D}_{\mathcal{P}}(\mathbf{x},\varepsilon)$ satisfies ELE if and only if there exist constant $\lambda\in[0,1]$ such that for all $\mathbf{x}\in\mathcal{X}^n$

$$\mathbf{d}(\mathbf{x},\varepsilon)=\frac{\varepsilon}{n}\mathbf{1}+(\mathbf{x}-\mu_x\mathbf{1})\left\{\left(1+\frac{\varepsilon}{n\mu_x}\right)^{\lambda}-1\right\}. \tag{82}$$

Proof. *Necessity:* ELE implies B for $\alpha+\beta=1$. Therefore Lemma 7 holds implying that $\boldsymbol{\delta}(\mathbf{x})=\frac{1}{n}\left[\lambda\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x+\tilde{\theta}}+\mathbf{1}\right]$ for $\lambda\in[0,1]$. It follows that $\boldsymbol{\delta}(\tilde{\mathbf{x}})=\boldsymbol{\delta}(\mathbf{x})=\boldsymbol{\delta}(\hat{\mathbf{x}})$ is equivalent to either $\lambda=0$, or for $\lambda\in(0,1]$ requires that:

$$\frac{\hat{\mathbf{x}}-\hat{\mu}\mathbf{1}}{\hat{\mu}+\tilde{\theta}}=\frac{\mathbf{x}-\mu\mathbf{1}}{\mu+\tilde{\theta}}=\frac{\tilde{\mathbf{x}}-\tilde{\mu}\mathbf{1}}{\tilde{\mu}+\tilde{\theta}}. \tag{83}$$

Since $\tilde{\mathbf{x}}=\alpha\mathbf{x}+\beta\hat{\mathbf{x}}$, then $\tilde{\mu}=\alpha\mu+\beta\hat{\mu}$, it follows that

$$\tilde{\mathbf{x}}-\tilde{\mu}\mathbf{1}=\alpha\left[\mathbf{x}-\mu\mathbf{1}\right]+\beta\left[\hat{\mathbf{x}}-\hat{\mu}\mathbf{1}\right]. \tag{84}$$

Recalling that from the first part of (83) $\hat{\mathbf{x}}-\hat{\mu}\mathbf{1}=\frac{\hat{\mu}+\tilde{\theta}}{\mu+\tilde{\theta}}\left[\mathbf{x}-\mu\mathbf{1}\right]$, after substituting into (84) and dividing by $\tilde{\mu}+\tilde{\theta}>0$ we get

$$\frac{\tilde{\mathbf{x}}-\tilde{\mu}\mathbf{1}}{\tilde{\mu}+\tilde{\theta}}=\left(\frac{\alpha(\mu+\tilde{\theta})+\beta(\hat{\mu}+\tilde{\theta})}{\tilde{\mu}+\tilde{\theta}}\right)\frac{\mathbf{x}-\mu\mathbf{1}}{\mu+\tilde{\theta}}.$$

Comparing with the second part of (83) we obtain the restriction $\alpha(\mu+\tilde{\theta})+\beta(\hat{\mu}+\tilde{\theta}) = \tilde{\mu}+\tilde{\theta}$ for all $\alpha > 0, \beta > 0$, and all $\mu \neq \hat{\mu}$. Recalling that $\tilde{\mu} = \alpha\mu+\beta\hat{\mu}$ we get $\alpha\tilde{\theta}+\beta\tilde{\theta} = \tilde{\theta}$ for all $\alpha > 0, \beta > 0$ implying $\tilde{\theta} = 0$. Substituting into the result in Proposition 3 gives the solution.

Sufficiency: Check that (82) is in $\mathcal{D}_{\mathcal{P}}(\mathbf{x}, \varepsilon)$ and given that the associated $\boldsymbol{\delta}(\mathbf{x})$ is $\boldsymbol{\delta}(\mathbf{x}) = \frac{1}{n}\left[\lambda\frac{\mathbf{x}-\mu_x\mathbf{1}}{\mu_x} + \mathbf{1}\right]$ satisfies ELE. ■

References

[1] Aczél J. (1966): *Lectures on Functional Equations and Their Applications.* New York, London: Academic Press.

[2] Aczél J. (1987): *A Short Course on Functional Equations.* Dordrecht: Reidel Publishing Co.

[3] Aczél J. and Dhombres, J. (1989): *Functional Equations in Several Variables.* Cambridge: Cambridge University Press.

[4] Aczél J. and Moszner, Z. (1994): New results on "scale" and "size" arguments justifying invariance properties of empirical indices and laws. *Mathematical Social Sciences,* **28**, 3–33.

[5] Amiel, Y. and Cowell, F. (1992): Measurement of income inequality. Experimental text by questionnaire. *Journal of Public Economics*, **47**, 3-26.

[6] Amiel, Y. and Cowell, F. (1997): Income transformation and income inequality. STICERD, LSE discussion paper DARP 24.

[7] Amiel, Y. and Cowell, F. (1999): *Thinking About Inequality.* Cambridge: Cambridge University Press.

[8] Atkinson, A. B. (1970): On the measurement of inequality. *Journal of Economic Theory*, **2**, 244-263.

[9] Atkinson, A. B. (1983): *Social Justice and Public Policy.* Wheatsheaf Books ltd.: UK.

[10] Aumann, R. J. and Maschler, M. (1985): Game theoretic analysis of a bankruptcy from the talmud. *Journal of Economic Theory*, **36**, 195-213.

[11] Ballano, C. and Ruiz-Castillo, J. (1993): Searching by questionnaire for the meaning of income inequality. *Revista Española de Economia,* **10**, 233-259.

[12] Berge (1963): *Topological Spaces.* Edinburgh and London: Oliver & Boid.

[13] Berrebi Z. M. and Silber, J. (1985): The Gini coefficient and negative incomes: a comment. *Oxford Economic Papers, **37**, 525-526.*

[14] Besley, T. P. and Preston, I. P. (1988): Invariance and the axiomatics of income tax progression: a comment. *Bulletin of Economic Research,* **40**, 159-163.

[15] Blackorby, C. and Donaldson, D. (1984): Ethically significant ordinal index of relative inequality. In Advances in Econometrics, Vol. 3, (Basman, R. L. and Rhodes, G. F. ed.), 131-47. London: JAI Press inc.

[16] Bosi, G., Candeal, J.C. and Indurain, E. (2000). Continuous representability of homothetic preferences by means of homogeneous utility functions. *Journal of Mathematical Economics,* **33**, 291-298.

[17] Bossert, W. and Pfingsten, A. (1990): Intermediate inequality: concepts, indices and welfare implications. *Mathematical Social Sciences,* **19**, 117-134.

[18] Bossert, W. (1998): Comment on 'The empirical acceptance of compensation axioms'. In J.-F. Laslier, M. Fleurbaey, N. Gravel, and A. Trannoy (eds.), Freedom in Economics: New Perspectives in Normative Analysis, Routledge, London, pp. 282-284.

[19] Candeal, J.C. and Indurain, E. (1993). On the structure of homothetic functions. *Aequationes Mathematicae*, **45**, 207-218.

[20] Candeal, J.C. and Indurain, E. (1995): Homothetic and weakly homothetic preferences. *Journal of Mathematical Economics,* **24**, 147-158.

[21] Chakravarty, S. R. (1988): On quasi orderings of income profiles. University of Paderborn, *Methods of Operations Research 60*, XIII Symposium on Operations Research, 455-473.

[22] Chen, C., Tsaur, T. and Rhai, T. (1982): The Gini coefficient and negative incomes. *Oxford Economic Papers,* ***34****,* 473-478. Reply (1985), ***37****, 52*7-528.

[23] Cowell, F. A. (2000): Measurement of Inequality. In *Handbook of Distribution* (Atkinson, A. B. and Bourguignon, F. eds.) North Holland.

[24] Chun, Y. (1988): The proportional solution for rights problems. *Mathematical Social Sciences,* **15**, 231-246.

[25] Dalton, H. (1920): The measurement of the inequality of incomes. *Economic Journal*, **20**, 348-361.

[26] Dasgupta, P., Sen, A. K. and Starrett, D. (1973): Notes on the measurement of inequality. *Journal of Economic Theory*, **6**, 180-187.

[27] Del Rio, C and Ruiz-Castillo, J. (2000): Intermediate inequality and welfare. *Social Choice and Welfare,* **17**, 223-239.

[28] Dutta, B. and Esteban, J. (1992): Social welfare and equality. *Social Choice and Welfare*, **9**, 267-276.

[29] Ebert, U. (1987): Size and distribution of income as determinants of social welfare. *Journal of Economic Theory*, **41**, 23-33..

[30] Ebert, U. (2000): Equivalizing incomes: A normative approach. *International Tax and Public Finance,* **7**, 619-640.

[31] Ebert, U. (2004): Coherent inequality views: linear invariant measures reconsidered. *Mathematical Social Sciences*, **47**, 1-20.

[32] Ebert, U. and Moyes, P. (2002): Equivalence scales reconsidered, forthcoming in *Econometrica.*

[33] Eichhorn, W. (1978): *Functional Equations in Economics.* Reading: Addison-Wesley.

[34] Eichhorn, W. and Gehrig, W. (1982): Measurement of inequality in economics. In *Optimization and Operations Research* (Korte, B. ed.) 657-693. Amsterdam: North Holland.

[35] Fields, G. S. and Fei, C. H. (1978): On inequality comparisons. *Econometrica,* **46**, 303-316.

[36] Foster, J. (1985): Inequality measurement. In *Fair Allocation* (Young, H. P. ed.)*;* Proceeding of Symposia in Applied Mathematics **33,** 31-68. Providence: The American Mathematical Society.

[37] Gelbaum, B. R. and Olmsted, J.M.H. (1964): *Counterexamples in analysis.* S. Francisco: Holden-Day Inc.

[38] Hardy, G.H., Littlewood, J.E., and Polya, G. (1934), *Inequalities,* London, Cambridge University Press.

[39] Harrison, E. and Seidl, C. (1994a):Perceptional inequality and preferential judgements: an empirical examination of distributional axioms. *Public Choice,* **79**, 61-81.

[40] Harrison, E. and Seidl, C. (1994b): Acceptance of distributional axioms: experimental findings. In *Models and Measurement of Welfare and Inequality* (Eichhorn, W. ed.) pp 67-99. Springer Verlag: Berlin.

[41] Kahneman, D. and Tversky, A. (1978): Prospect theory: an analysis of decision under risk. *Econometrica*, **47**, 263-291.

[42] Krtscha (1994): A new compromise measure of inequality. In *Models and Measurement of Welfare and Inequality* (Eichhorn, W. ed.). Springer Verlag: Berlin.

[43] Kolm, S. C. (1969): The optimal production of social justice. In *Public Economics* (Margolis, J. e Gutton, H. ed.), pp. 145-200. London: Mcmillan.

[44] Kolm, S. C. (1976a): Unequal inequalities. I. *Journal of Economic Theory*, **12**, 416-442.

[45] Kolm, S. C. (1976b): Unequal inequalities. II. *Journal of Economic Theory*, **13**, 82-111.

[46] Kolm, S. C. (1996): Intermediate measures of inequality. Mimeo.

[47] Lambert, P. J. (2001): *The Distribution and Redistribution of Income: a Mathematical Analysis.* 2nd edition, Manchester: Manchester University Press.

[48] Marshall, A. W. and Olkin, I. (1979): *Inequalities: Theory of Majorization and Its Applications.* New York: Academic Press.

[49] Moulin, H. (1985a): Egalitarianism and utilitarianism in quasi linear bargaining. *Econometrica*, **53**, 49-68.

[50] Moulin, H. (1985b): The separability axiom and equal sharing methods. *Journal of Economic Theory*, **36**, 120-148.

[51] Moulin, H. (1987): Equal or proportional division of a surplus, and other methods. *International Journal of Game Theory,* **16**, 161-186.

[52] Moulin, H. (2000): Priority rules and other asymmetric rationing methods. *Econometrica*, **68**, 643-684.

[53] Moulin, H. (2002): Axiomatic cost and surplus sharing. In *Handbook of Social Choice and Welfare*, Arrow, K. Sen, A. K. and Suzumura, K eds, North-Holland.

[54] Pfingsten, A. (1986): Distributionally neutral tax changes for different inequality concepts. *Journal of Public Economics.* **30**, 385-393.

[55] Pfingsten, A. (1991): Surplus sharing methods. *Mathematical Social Sciences,* **21**, 287-301.

[56] Pigou, A. C. (1912): *Wealth and Welfare.* MacMillan: New York.

[57] Rothschild, M. and Stiglitz, J. E. (1973): Some further results on the measurement of inequality. *Journal of Economic Theory*, **6**, 188-204.

[58] Seidl, C. and Pfingsten, A. (1997): Ray invariant inequality measures. In *Research on Economic Inequality,* Vol. 7 pp 107-129. JAI Press.

[59] Sen, A. K. (1973): *On Economic Inequality.* Oxford: Claredon Press. (1997) expanded edition with the annexe "*On Economic Inequality After a Quarter Century*" by Foster, J. And Sen, A.K.

[60] Wakker, P. (1987): From decision making under uncertainty to game theory. In *Surveys in Game Theory and Related Topics* (Peters, H. and Vrieze, O. J. eds.) 163-180. CWI Tract 39, Amsterdam.

[61] Yoshida, T. (2002): Social welfare rankings of income distributions. A new parametric concept of intermediate inequality. Discussion Paper I-46, Department of Economics, Okayama University.

[62] Young, H. P. (1987): On dividing an amount according to individual claims or liabilities. *Mathematics of Operations Research,* **12**, 398-414.

[63] Young, H. P. (1987a): Progressive taxation and the equal sacrifice principle. *Journal of Public Economics.* **32**, 203-214.

[64] Young, H. P. (1988):Distributive justice in taxation. *Journal of Economic Theory*, **44**, 321-335.

[65] Young, H. P. (1988a): Equal sacrifice in taxation. In *Measurement in Economics* (Eichhorn, W. ed.), 563-574. Heidelberg: Physica Verlag.

[66] Zheng, B. (1994): Can a poverty index be both relative and absolute? *Econometrica,* **62**, 1453-1458.

[67] Zheng, B. (2002): On intermediate measures of inequality. Mimeo, Department of Economics, University of Colorado.

[68] Zheng, B. (2003): Unit-consistent decomposable inequality measures. Mimeo, Department of Economics, University of Colorado.

[69] Zheng, B. (2004): Unit-consistent poverty indices. Mimeo, Department of Economics, University of Colorado.

[70] Zoli, C. (1998): A surplus sharing approach to the measurement of inequality. University of York, Discussion Paper 98/25.

[71] Zoli, C. (1999): A generalized version of the inequality equivalence concept, with applications to the measurement of inequality. In *Logic, Game Theory and Social Choice* (De Swart ed.) 427-441. Tilburg University Press: Tilburg.

[72] Zoli, C. (2002): Inequality, Welfare and Poverty Comparisons. Ph.D. Dissertation, Department of Economics and Related Studies, University of York.

Annex

Detailed derivation of proofs in §3.3 and §3.4.

We first characterize IEDVs satisfying LC and VE.

Lemma 8 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies LC and VE if and only if it is the same as obtained in Lemma 1 and in addition satisfies the condition:*

$$\left[\varepsilon - \sum\nolimits_{h=1}^{n} a_h(\mu,\varepsilon)\right] \geq -\min{}_j \left\{\left[a_j(\mu,\varepsilon) - a_{j-1}(\mu,\varepsilon)\right] \cdot (n-j+1) : j \in \mathcal{N}\right\} \quad (85)$$

for all $\mu,\varepsilon > 0$*, where* $x_i \geq x_{i-1}$ *for all* $i \in \mathcal{N}$ *and* $a_0(\mu,\varepsilon) = 0$.

Proof of Lemma 8:

Given Lemma 1, applying VE: $d_j(\mathbf{x},\varepsilon) \geq d_i(\mathbf{x},\varepsilon)$ for all $\mathbf{x} \in X^n$ and all $\mu,\varepsilon > 0$ if $x_j > x_i$, we get

$$a_j(\mu,\varepsilon) + \frac{\varepsilon - \sum_{h=1}^{n} a_h(\mu,\varepsilon)}{n\mu} x_j \geq a_i(\mu,\varepsilon) + \frac{\varepsilon - \sum_{h=1}^{n} a_h(\mu,\varepsilon)}{n\mu} x_i$$

for all $j, i \in \mathcal{N}$ such that $x_j > x_i$, all $\mathbf{x} \in X^n$ and all $\mu,\varepsilon > 0$. Rearranging we get

$$\left[\varepsilon - \sum\nolimits_{h=1}^{n} a_h(\mu,\varepsilon)\right] \geq -\left[a_j(\mu,\varepsilon) - a_i(\mu,\varepsilon)\right] \frac{n\mu}{(x_j - x_i)}. \quad (86)$$

Note that, since $a_j(\mu,\varepsilon) \geq a_i(\mu,\varepsilon)$, then the r.h.s. of (86) is non-positive. Consider the set of ordered distributions $\mathcal{X}^n$ where $x_i \geq x_{i-1}$ for all $i \in \mathcal{N}$. Since (86) has to be satisfied for all $\mathbf{x} \in X^n$ and the l.h.s. is the same for all distributions $\mathbf{x} \in \mathcal{X}^n$, then it will be sufficient to check that the condition is satisfied given the smallest value of $\left[a_j(\mu,\varepsilon) - a_i(\mu,\varepsilon)\right] / (x_j - x_i)$. We first consider the largest value of $(x_j - x_i)$. This value can be obtained for the ordered distribution $\mathbf{x} \in \mathcal{X}^n(\mu)$ such that $x_j = x_{j+1} = ... = x_n = \frac{n\mu}{n-j+1}$ and $x_1 = x_2 = .. = x_i = x_{j-1} = 0$. It follows that $\max(x_j - x_i) = \frac{n\mu}{n-j+1}$. Since the largest value of $(x_j - x_i)$ is independent from i we can also consider taking the minimum value of $a_j(\mu,\varepsilon) - a_i(\mu,\varepsilon)$. Given that $a_j(\mu,\varepsilon) \geq a_i(\mu,\varepsilon)$ for all $x_j \geq x_i$, then $\min\left[a_j(\mu,\varepsilon) - a_i(\mu,\varepsilon)\right] = a_j(\mu,\varepsilon) - a_{j-1}(\mu,\varepsilon)$ for all $j \in \mathcal{N}$ where by definition is set $a_0(\mu,\varepsilon) = 0$ in order to include the value of $j = 1$.

After substituting and recalling that if (86) is satisfied for the highest value on the r.h.s. it is satisfied for all the possible values obtained on the r.h.s. we get:

$$\left[\varepsilon - \sum\nolimits_{h=1}^{n} a_h(\mu,\varepsilon)\right] \geq -\min{}_j \left\{\left[a_j(\mu,\varepsilon) - a_{j-1}(\mu,\varepsilon)\right](n-j+1) : j \in \mathcal{N}\right\}$$

for all $\mu,\varepsilon > 0$. ■

It follows that $\varepsilon \geq \sum_{h=1}^{n} a_h(\mu,\varepsilon)$ is sufficient to guarantee VE, but in general we may also have $\varepsilon + n\mu \geq \sum_{h=1}^{n} a_h(\mu,\varepsilon) > \varepsilon$.

Note that if HE is added then $a_j(\mu,\varepsilon) = a_i(\mu,\varepsilon)$ for all $j, i \in \mathcal{N}$, and all $\mu,\varepsilon > 0$. It follows that $-\min_j \left\{\left[a_j(\mu,\varepsilon) - a_i(\mu,\varepsilon)\right](n-j+1) : j, i \in \mathcal{N},\ j > i\right\} = 0$ giving $\varepsilon \geq \sum_{h=1}^{n} a_h(\mu,\varepsilon)$.

5.1.7 Proposition 7

First we prove the following Lemma and Remark 4.

Lemma 9 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies conditions LC and ADD if and only if there exists a set of continuous functions* $\psi_i : \mathbb{R}_{++} \to [0,1]$ *such that (i')* $\psi_i(\mu) \geq \psi_{i-1}(\mu)$ *for all* $\mu > 0$, $i \in \mathcal{N}$ *, (ii')* $1 \geq \sum_{i=1}^n \psi_i(\mu) \geq 0$ *for all* $\mu > 0$, *and for all* $\mathbf{x} \in \mathcal{X}^n$

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[n\psi_i(\mu_x) + (1-\Psi(\mu_x))\frac{x_i}{\mu_x}\right] \quad \forall i \in \mathcal{N}, \tag{87}$$

where $\Psi(\mu_x) = \sum_{i=1}^n \psi_i(\mu_x)$.

Proof of Lemma 9:

Consider ADD: $\mathbf{d}(\mathbf{x},\varepsilon) + \mathbf{d}(\mathbf{x},\varepsilon') = \mathbf{d}(\mathbf{x},\varepsilon+\varepsilon')$ for all $\varepsilon, \varepsilon' > 0$, all $\mathbf{x} \in X^n$. After repeated application of ADD we get $\sum_{j=1}^{\lambda} \mathbf{d}(\mathbf{x},\varepsilon_i) = \mathbf{d}(\mathbf{x},\sum_{j=1}^{\lambda}\varepsilon_j)$ for all $\lambda \in \mathbb{N}, \varepsilon_j > 0$ all $\mathbf{x} \in X^n$. Suppose $\varepsilon_j = \varepsilon \; \forall j = 1,2,..,\lambda$, then we obtain $\lambda\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\mathbf{x},\lambda\varepsilon)$ for all $\lambda \in \mathbb{N}$, $\varepsilon > 0$ all $\mathbf{x} \in X^n$. From Lemma 1 we know that for all IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfying condition LC : $d_i(\mathbf{x},\varepsilon) = a_i(\mu,\varepsilon) + \frac{\varepsilon - \sum_{j=1}^n a_j(\mu,\varepsilon)}{n\mu} x_i \; \forall i \in \mathcal{N}$. Substituting into $\lambda\mathbf{d}(\mathbf{x},\varepsilon) = \mathbf{d}(\mathbf{x},\lambda\varepsilon)$ for all $\lambda \in \mathbb{N}$ we obtain:

$$\lambda a_i(\mu,\varepsilon) + \frac{\lambda\varepsilon - \lambda\sum_{j=1}^n a_j(\mu,\varepsilon)}{n\mu}x_i = a_i(\mu,\lambda\varepsilon) + \frac{\lambda\varepsilon - \sum_{j=1}^n a_j(\mu,\lambda\varepsilon)}{n\mu}x_i$$

that is

$$\lambda a_i(\mu,\varepsilon) - \left[\sum\nolimits_{j=1}^n \lambda a_j(\mu,\varepsilon)\right]\frac{x_i}{n\mu} = a_i(\mu,\lambda\varepsilon) - \left[\sum\nolimits_{j=1}^n a_j(\mu,\lambda\varepsilon)\right]\frac{x_i}{n\mu} \quad \forall i \in \mathcal{N} \tag{88}$$

and for all $\lambda \in \mathbb{N}$, $\varepsilon, \mu > 0$. The functional equation in (88) is satisfied if and only if:

$$\lambda a_i(\mu,\varepsilon) = a_i(\mu,\lambda\varepsilon) \quad \forall\mu > 0,\ \varepsilon > 0,\ i \in \mathcal{N},\ \lambda \in \mathbb{N}. \tag{89}$$

Consider $\varepsilon, \varepsilon' > 0$ and $\lambda, \lambda' \in \mathbb{N}$ such that $\lambda\varepsilon = \lambda'\varepsilon' = k$. From (89) follows that $\lambda a_i(\mu,\varepsilon) = a_i(\mu,k) = \lambda' a_i(\mu,\varepsilon')$ for all $i \in \mathcal{N}$ all $\mu > 0$. Letting $\rho := \lambda/\lambda'$ and noting that $\varepsilon' = \lambda\varepsilon/\lambda' = \rho\varepsilon$ it follows that (89) is equivalent to

$$\rho a_i(\mu,\varepsilon) = a_i(\mu,\rho\varepsilon) \quad \forall\mu > 0,\ \varepsilon > 0,\ i \in \mathcal{N},\ \rho \in \mathbb{Z}_{++}, \tag{90}$$

where $\mathbb{Z}_{++}$ is the set of positive rational numbers. The general solution of (90) where $a_i(\mu,\varepsilon)$ is continuous, is obtained letting $\varepsilon = 1$, it follows that letting $\psi_i(\mu) := a_i(\mu,1)$ for all $\mu > 0$

$$a_i(\mu,\rho) = \rho \cdot a_i(\mu,1) = \psi_i(\mu)\cdot\rho \quad \forall\mu > 0,\ i \in \mathcal{N},\ \rho \in \mathbb{Z}_{++}. \tag{91}$$

Given continuity of $a_i(\mu,\varepsilon)$ w.r.t. ε, since the set of positive rational numbers is dense, then the result can be extended to the closure of the set i.e. $a_i(\mu,\varepsilon) = \psi_i(\mu)\cdot\varepsilon$ for all $\mu, \varepsilon > 0$, all $i \in \mathcal{N}$.

Following Lemma 1 we have: (i') $\psi_j(\mu) \geq \psi_i(\mu) \geq 0$ if $x_j \geq x_i$, (ii) $1 + \min_i \{\psi_i(\mu)\cdot(n-i+1) : i \in \mathcal{N}\} \geq \sum_{i=1}^n \psi_i(\mu) \geq 0$ if $x_n \geq x_i$ for all $i \in \mathcal{N}$, and

(iii) $\varepsilon + n\mu \geq \varepsilon \cdot \sum_{i=1}^{n} \psi_i(\mu)$, for all μ,$\varepsilon > 0$. Rearranging condition (iii) dividing both sides by ε we get $1 + \frac{n\mu}{\varepsilon} \geq \sum_{i=1}^{n} \psi_i(\mu)$. Since the condition has to hold for all $\varepsilon > 0$, then letting $\varepsilon \to \infty$ we obtain $1 \geq \sum_{i=1}^{n} \psi_i(\mu)$ that also implies that condition (ii) is satisfied given that $\psi_i(\mu) \geq 0$.

As a result the functions $\psi_i(\mu)$ has to satisfy the conditions (i') $\psi_j(\mu) \geq \psi_i(\mu) \geq 0$ if $x_j \geq x_i$, and (ii') $1 \geq \sum_{i=1}^{n} \psi_i(\mu) \geq 0$. Letting $\Psi(\mu_x) = \sum_{i=1}^{n} \psi_i(\mu_x)$, and substituting into (29) we get (87) . ■

Remark 4:For all IEDV's $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x}, \varepsilon)$:

(i) $ADD \Leftrightarrow LE$.

(ii) $LC + ADD \Longrightarrow VE$.

(iii) $LC + PI + WCIC \Longrightarrow VE$.

Proof of Remark 4:

Part (i) Let $\varepsilon' = \varepsilon$ in ADD. After repeated application of ADD we get that for all $r, t \in \mathbb{N}$, $r\mathbf{d}(\mathbf{x}, \varepsilon) = \mathbf{d}(\mathbf{x}, r\varepsilon) = t\mathbf{d}(\mathbf{x}, r\varepsilon/t) \Leftrightarrow r\varepsilon\left[\mathbf{d}(\mathbf{x}, \varepsilon)/\varepsilon\right] = r\varepsilon\left[\mathbf{d}(\mathbf{x}, k\varepsilon)/k\varepsilon\right]$ where $k = r/t$. That is $\mathbf{d}(\mathbf{x}, \varepsilon)/\varepsilon = \mathbf{d}(\mathbf{x}, k\varepsilon)/k\varepsilon$ for all $\varepsilon > 0$, all $k \in \mathbb{Z}_{++}$ (i.e. all positive rational numbers k). Given continuity of $\mathbf{d}(\mathbf{x}, \varepsilon)$ w.r.t. ε, the last condition is equivalent to LE i.e. $\mathbf{d}(\mathbf{x}, \varepsilon)/\varepsilon = \mathbf{d}(\mathbf{x}, \varepsilon')/\varepsilon'$ for all $\varepsilon, \varepsilon' > 0$.

Part (ii) From (87) in Lemma 9 is clear that the IEDVs satisfy VE $d_i(\mathbf{x},\varepsilon) \geq d_{i-1}(\mathbf{x},\varepsilon)$ for all $i \in \mathcal{N}$, all $\varepsilon > 0$ all $\mathbf{x} \in X^n$, i.e. $n\left[\psi_i(\mu_x) - \psi_{i-1}(\mu_x)\right] + (1 - \Psi(\mu_x))\frac{x_i - x_{i-1}}{\mu_x} \geq 0$ for all $\mu_x > 0$. ■

Recall that, as shown in the proof of Proposition 1 HE is not sufficient together with LC to imply VE.

Part (iii): We first show that WCIC is not sufficient together with LC to imply VE, from Lemma 10 for instance if $\phi_i(t) = \phi(t) \in (1/n; \min\{\frac{1}{n-1}; t + \frac{1}{n}\}]$ then VE is not satisfied. In Proposition 14: we show that LC, PI and WCIC imply VE proving the remark.

Proposition 7 (Pfingsten, 1991): An IEC is characterized by IEDVs satisfying LC, HE, VE, PI, ADD if and only if it is: for all $\mathbf{x}, \mathbf{y} \in X^n$ if either

$$\frac{y_i - \mu_y}{\mu_y + \theta} = \frac{x_i - \mu_x}{\mu_x + \theta} \quad \forall i \in \mathcal{N} \tag{92}$$

where $\theta \geq 0$, or

$$y_i - \mu_y = x_i - \mu_x \quad \forall i \in \mathcal{N}$$

then $\mathbf{x} \sim_E \mathbf{y}$.

Proof of Proposition 7:

Necessity: Recall that for all $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ ADD is equivalent to LE that in turns is equivalent to $\gamma(\mu,\varepsilon) = \tilde{\gamma}(\mu)$ for all $\mu, \varepsilon > 0$, and that VE is implied by ADD+LC. Comparing the definitions of PI and ADD then we can redefine PI+ADD for all $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$ as equivalent to $\gamma(\mu,\varepsilon) = \tilde{\gamma}(\mu)$ and $\mathbf{d}(\mathbf{x}, \varepsilon') = \mathbf{d}\left[\mathbf{x} + \mathbf{d}(\mathbf{x}, \varepsilon), \varepsilon'\right]$ for all $\mathbf{x} \in X^n$, $\forall \varepsilon, \varepsilon' > 0$ for all $\mathbf{d}(\mathbf{x}, \varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x}, \varepsilon)$. Making use of the same notation as in Lemma 3 we can express this condition as $\tilde{\gamma}(\mu)\boldsymbol{\alpha} = \tilde{\gamma}(\mu+e)\boldsymbol{\beta}$ where $e := \varepsilon/n$. Rearranging (40) we get $e\tilde{\gamma}(\mu) + e'\tilde{\gamma}(\mu+e)\left[\mu + e\tilde{\gamma}(\mu)\right]/(\mu+e) = (e+e')\,\tilde{\gamma}(\mu)$ that is, after simplifications,

$$\tilde{\gamma}(\mu) = \tilde{\gamma}(\mu+e)\frac{\mu + e\tilde{\gamma}(\mu)}{\mu+e} \quad \text{for all } \mu, e > 0, \tag{93}$$

which can be rewritten as:

$$\mu\tilde{\gamma}(\mu) + e\tilde{\gamma}(\mu) = \mu\tilde{\gamma}(\mu+e) + e\tilde{\gamma}(\mu)\tilde{\gamma}(\mu+e).$$

Subtracting from both sides $\mu\tilde{\gamma}(\mu)\tilde{\gamma}(\mu+e)$ and rearranging we get:

$$(\mu+e)\tilde{\gamma}(\mu)\left[1-\tilde{\gamma}(\mu+e)\right] = \mu\tilde{\gamma}(\mu+e)\left[1-\tilde{\gamma}(\mu)\right] \quad \text{for all } \mu, e > 0. \tag{94}$$

A solution for this functional equation is $\tilde{\gamma}(\mu) = 0\ \forall\mu > 0$. If $\tilde{\gamma}(\hat{\mu}) = 0$ for some $\hat{\mu} > 0$, then according to (94) letting $\mu = \hat{\mu}$ we get $0 = \hat{\mu}\tilde{\gamma}(\hat{\mu}+e)$ for all $e > 0$, that is $\tilde{\gamma}(\mu) = 0$ for all $\mu > \hat{\mu}$, while letting $\mu + e = \hat{\mu}$ we get $\hat{\mu}\tilde{\gamma}(\hat{\mu}-e) = 0$ for all $\hat{\mu} > e > 0$ that is $\tilde{\gamma}(\mu) = 0$ for all positive $\mu < \hat{\mu}$. It follows that according to (94) $\tilde{\gamma}(\mu) = 0$ for some $\mu > 0$ implies that $\tilde{\gamma}(\mu) = 0\ \forall\mu > 0$.

On the other hand if $\tilde{\gamma}(\mu) \neq 0\ \forall\mu > 0$, then we can rewrite (94) as:

$$(\mu+e)\frac{1-\tilde{\gamma}(\mu+e)}{\tilde{\gamma}(\mu+e)} = \mu\frac{1-\tilde{\gamma}(\mu)}{\tilde{\gamma}(\mu)}\ \forall\mu, e > 0.$$

Letting $\varphi(\mu) = \mu\frac{1-\tilde{\gamma}(\mu)}{\tilde{\gamma}(\mu)}$, the previous functional equation can be rewritten as:

$$\varphi(\mu) = \varphi(\mu+e)\ \forall\mu, e > 0$$

where, since $\tilde{\gamma}(\mu) \in (0,1]$, then $\varphi(\mu) > 0\ \forall\mu > 0$. The previous functional equation is equivalent to $\varphi(\mu)$ being constant, that is there exists some real number $\tilde{\theta} > 0$ such that $\varphi(\mu) = \tilde{\theta}$ for all $\mu > 0$. Substituting we get $\tilde{\gamma}(\mu) = \frac{\mu}{\mu+\tilde{\theta}}$ for all $\mu > 0$. Thus, the solution of (93) is

$$\text{either } \tilde{\gamma}(\mu) = 0 \quad \text{or} \quad \tilde{\gamma}(\mu) = \frac{\mu}{\mu+\tilde{\theta}} \quad \text{for all } \mu > 0. \tag{95}$$

giving the IIE criterion. Just substitute for either $\eta(\mu) = \tilde{\gamma}(\mu) = 0$ or $\eta(\mu) = \tilde{\gamma}(\mu) = \frac{\mu}{\mu+\theta}$ respectively from (95) into (5) and solve. ■

Sufficiency: Check that LC,HE, ADD and PI are satisfied.

5.1.8 Proposition 8

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC, HE, PI and WCIC if and only if there exists a constant $\lambda \in [0,1]$ such that for all $\mathbf{x} \in X^n$, and $\varepsilon, \varepsilon' > 0$:

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{1} + (\mathbf{x} - \mu_x\mathbf{1})\left\{\left(1 + \frac{\varepsilon}{n}\cdot\frac{1}{\mu_x}\right)^{\lambda} - 1\right\} \quad \lambda \in [0,1]\,. \tag{96}$$

Proof of Proposition 8:

Necessity: We apply HE to the result in Proposition 14. According to (126) if $x_i = x_{i-1}$ we get

$$d_i(\mathbf{x},\varepsilon) - d_{i-1}(\mathbf{x},\varepsilon) = (\upsilon_i - \upsilon_{i-1}) \cdot \mu_x \cdot \left[\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right) - \left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda} \right].$$

HE is therefore satisfied iff $d_i(\mathbf{x},\varepsilon) - d_{i-1}(\mathbf{x},\varepsilon) = 0$ implying that either $\lambda = 1$, or $\upsilon_i = 1$ for all $i \in \mathcal{N}$. Thereby transforming (126) into (96).

Sufficiency: Check that (96) satisfies HE, LC, PI and WCIC. ■

5.1.9 Proposition 9

For any IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ the following relations hold:

1. $ADD + B \rightarrow EB$,
2. $PI + ADD \rightarrow B$,
3. $PI + ADD \longleftrightarrow PI + EB$.
4. $PI + WCIC \longleftrightarrow PI + ELE$
5. $PI + WCIC + ADD \longleftrightarrow PI + EE$.

Proof of Proposition 9:

We need just to prove the first 2 statements:

Part 1*:* The result in Proposition 13 derived considering ADD+B (in addition to LC, HE, VE) can be obtained letting $\omega(\varepsilon) = \tilde{\omega}$ and $\theta(\varepsilon) = \tilde{\theta}$ in Proposition 4 where EB is considered.

Part 2*:* This solution is the same as that which could be obtained from the characterization of $\eta(\mu)$ associated with ADD+B, that is $\eta(\mu) = \lambda\frac{\mu}{\mu+\tilde{\theta}}$ when either $\lambda = 1$, $\tilde{\theta} \geq 0$, or $\lambda = 0$. On the other hand it is possible to show that conditions PI+ADD imply PI+B simply recalling that the restriction in (95) could be obtained from those shown in Proposition 3 for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfying conditions PI+B, again letting either $\lambda = 1$, $\tilde{\theta} \geq 0$, or $\lambda = 0$.

5.1.10 Proposition 10

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies conditions LC, ADD, PI if and only if there exist $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$, such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$ for any $i = 1, 2, ...n$, $\sum_{i=1}^{n} \upsilon_i = n$, and a constant $\beta \geq 0$ such that for all $\mathbf{x} \in \mathcal{X}^n$, $\varepsilon > 0$: either

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\frac{\upsilon_i\beta + x_i}{\beta + \mu_x}\right], \quad \forall i \in \mathcal{N}, \tag{97}$$

or

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\upsilon_i, \quad \forall i \in \mathcal{N}. \tag{98}$$

Proof of Proposition 10:

Sufficiency: Check that (97) and (98) satisfy conditions VE, ADD, PI, LC.

Necessity: Add PI to the result in Lemma 9. We know from Proposition 7 that $(ADD + PI) \longleftrightarrow \mathbf{d}(\mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon),\varepsilon') = \mathbf{d}(\mathbf{x},\varepsilon') + ADD$. Applying to (87) in Lemma 9, letting $e := \varepsilon/n$, denoting with $\mathbf{y} = \mathbf{x} + \mathbf{d}(\mathbf{x},\varepsilon)$ and recalling that $\mu_y = \mu_x + e$ we obtain:

$$e'\left[n\psi_i(\mu_x) + (1 - \Psi(\mu_x))\frac{x_i}{\mu_x}\right] = e'\left[n\psi_i(\mu_x + e) + (1 - \Psi(\mu_x + e))\frac{y_i}{\mu_x + e}\right] \quad \forall i \in \mathcal{N} \tag{99}$$

and for all $e', e, \mu_x > 0$. According to the definition of $\mathbf{y}$:

$$y_i = x_i + e\left[n\psi_i(\mu_x) + (1 - \Psi(\mu_x))\frac{x_i}{\mu_x}\right] \quad \forall i \in \mathcal{N}.$$

Note that $\frac{y_i}{\mu_x+e} = \frac{x_i}{\mu_x}\frac{\mu_x+e(1-\Psi(\mu_x))}{\mu_x+e} + n\frac{e}{\mu_x+e}\psi_i(\mu_x)$, substituting into (99) and rearranging we get:

$$\begin{aligned} n\left[\psi_i(\mu_x) - \psi_i(\mu_x + e)\right] + \left[1 - \Psi(\mu_x)\right]\frac{x_i}{\mu_x} &= \left\{\left[1 - \Psi(\mu_x + e)\right]\left[1 - \frac{e\Psi(\mu_x)}{\mu_x + e}\right]\right\}\frac{x_i}{\mu_x} \\ +n\left[1 - \Psi(\mu_x + e)\right]\frac{e\psi_i(\mu_x)}{\mu_x + e} \qquad \forall i &\in \mathcal{N},\ \forall \mu_x > 0, e > 0. \end{aligned} \tag{100}$$

Both sides of the equation could be split into two parts, the first which does not depend on x_i, given a μ_x, and the second which does. In order to satisfy the previous equation for all $\mathbf{x} \in \mathcal{X}^n$ both parts on the two sides of the equation should equate. Suppose that $\tilde{\psi}_i(\mu_x,\varepsilon)$ satisfies (100) for a given distribution $\tilde{\mathbf{x}}$, it is always possible to find another distribution $\tilde{\mathbf{x}}' \neq \tilde{\mathbf{x}}$ with the same average income as $\tilde{\mathbf{x}}$, such that the solution $\tilde{\psi}_i(\mu_x,\varepsilon)$ is no longer valid except in the case in which the two multiplicative terms of x_i in both sides of (100) are satisfied for $\tilde{\Psi}(\mu_x,\varepsilon) = \sum_i \tilde{\psi}_i(\mu_x,\varepsilon)$. On the other hand, consider $x_i = 0$ for some $i \in \mathcal{N}$, in this case the first parts of each sides of (100) should equate. Thus:

$$\begin{cases} 1 - \Psi(\mu_x) = \left[1 - \Psi(\mu_x + e)\right]\left[1 - \frac{e\Psi(\mu_x)}{\mu_x+e}\right] \\ \psi_i(\mu_x) - \psi_i(\mu_x + e) = \left[1 - \Psi(\mu_x + e)\right]\frac{e\psi_i(\mu_x)}{\mu_x+e} \end{cases} \quad \forall i \in \mathcal{N},\ \forall \mu_x > 0, e > 0. \tag{101}$$

If we consider w.l.o.g. $e = (\lambda - 1)\mu_x$, $\lambda > 1$, the previous conditions become:

$$\begin{cases} 1 - \Psi(\mu_x) = \left[1 - \Psi(\lambda\mu_x)\right]\left[1 - \frac{(\lambda-1)}{\lambda}\Psi(\mu_x)\right] \\ \psi_i(\mu_x) - \psi_i(\lambda\mu_x) = \left[1 - \Psi(\lambda\mu_x)\right]\frac{(\lambda-1)}{\lambda}\psi_i(\mu_x) \end{cases} \quad \forall i \in \mathcal{N}, \mu_x > 0, \lambda > 1. \tag{102}$$

Rearranging the first condition we obtain:

$$1 - \Psi(\mu_x) = \frac{1}{\lambda}\left[1 - \Psi(\lambda\mu_x)\right]\left[\lambda\left(1 - \Psi(\mu_x)\right) + \Psi(\mu_x)\right] \quad \forall \mu_x > 0, \lambda > 1 \tag{103}$$

that is

$$\lambda\left[1 - \Psi(\mu_x)\right]\Psi(\lambda\mu_x) = \left[1 - \Psi(\lambda\mu_x)\right]\Psi(\mu_x) \quad \forall \mu_x > 0, \lambda > 1. \tag{104}$$

A solution of this functional equation is obtained for $\Psi(\mu_x) = 1$ for all $\mu_x > 0$. If $\Psi(\mu_x) = 1$ for some $\mu_x = \hat{\mu} > 0$ then according to (104) letting $\mu_x = \hat{\mu}$ we get $0 = 1 - \Psi(\lambda\hat{\mu}_x)$ for all $\lambda > 1$, that is $\Psi(\mu) = 1$ for all $\mu > \hat{\mu}$, while letting $\lambda\mu_x = \hat{\mu}$ we get $\lambda\left[1 - \Psi(\hat{\mu}/\lambda)\right] = 0$ for all $\lambda > 1$ that is $\Psi(\mu) = 1$ for all positive $\mu < \hat{\mu}$. It follows that according to (104) $\Psi(\mu_x) = 1$ for some $\mu_x > 0$ implies that $\Psi(\mu_x) = 1$ for all $\mu_x > 0$.

On the other hand if $\Psi(\mu_x) \neq 1$ for all $\mu_x > 0$ the previous condition could be rewritten as

$$\frac{\Psi(\mu_x)}{1 - \Psi(\mu_x)} = \frac{\Psi(\lambda\mu_x)}{1 - \Psi(\lambda\mu_x)}\lambda \qquad \forall \mu_x > 0, \lambda > 1. \tag{105}$$

Denoting with $\omega(\mu_x) = \frac{\Psi(\mu_x)}{1-\Psi(\mu_x)}$, we obtain:

$$\omega(\mu_x) = \omega(\lambda\mu_x)\lambda \quad \forall \mu_x > 0, \lambda > 1, \tag{106}$$

where, given the definition of $\Psi(\mu_x)$, $\omega(\mu_x) \geq 0$. The solution of this functional equation is:

$$\omega(\mu_x) = \frac{\beta}{\mu_x}, \text{ where } \beta \geq 0. \tag{107}$$

In order to prove this result, fix $\mu_x = 1$ in (106), then for all $\lambda > 1$ we obtain $\omega(\lambda) = \frac{\omega(1)}{\lambda}$. Let $\omega(1) = \beta \geq 0$, then $\omega(\mu) = \frac{\beta}{\mu}$ for all $\mu \geq 1$. The solution for all $1 > \mu > 0$ could be obtained fixing $\lambda\mu_x = 1$, substituting for $\mu_x = \frac{1}{\lambda}$, as $\lambda > 1$, we can obtain all values of $\mu \in (0, 1)$, that is $\omega(\frac{1}{\lambda}) = \frac{\omega(1)}{(1/\lambda)} = \frac{\beta}{(1/\lambda)}$ which gives $\omega(\mu) = \frac{\beta}{\mu}$ for all $\mu \in (0, 1)$.

Substituting for the definition of $\omega(\mu_x)$ we obtain $\frac{\Psi(\mu_x)}{1-\Psi(\mu_x)} = \frac{\beta}{\mu_x}$, that is:

$$\text{either } \Psi(\mu_x) = \frac{\beta}{\mu_x + \beta}, \text{ or } \Psi(\mu_x) = 1, \quad \forall \mu_x > 0 \tag{108}$$

where $\beta \geq 0$. Consider now the second functional relation in (102). Substituting for $\Psi(\mu_x) = \frac{\beta}{\mu_x+\beta}$ we obtain

$$\psi_i(\mu_x) - \psi_i(\lambda\mu_x) = \left[\frac{\lambda\mu_x}{\lambda\mu_x + \beta}\right] \frac{(\lambda - 1)}{\lambda}\psi_i(\mu_x), \quad \forall \mu_x > 0, \lambda > 1, i \in \mathcal{N}, \tag{109}$$

rearranging, the condition becomes, $\psi_i(\mu_x)\left[1 - \frac{\mu_x(\lambda-1)}{\lambda\mu_x+\beta}\right] = \psi_i(\lambda\mu_x)$, that is:

$$\psi_i(\mu_x)\,(\mu_x + \beta) = (\lambda\mu_x + \beta)\,\psi_i(\lambda\mu_x) \quad \forall \mu_x > 0, \lambda > 1, i \in \mathcal{N}. \tag{110}$$

The solution for this family of functional equations is:

$$\psi_i(\mu_x) = \frac{\varsigma_i}{\mu_x + \beta}, \text{ where } \varsigma_i \geq 0 \text{ and } \beta \geq 0, \ \forall \mu_x > 0, i \in \mathcal{N}.$$

Let $\kappa_i(\mu_x, \beta) := (\mu_x + \beta)\,\psi_i(\mu_x)$ where $\kappa_i(\mu_x, \beta)$ is a parametrized function for β, given its definition it is $\kappa_i(\mu_x, \beta) \geq 0$ for all $\mu_x > 0, \beta \geq 0$. Then the functional equation could be rewritten as $\kappa_i(\mu_x, \beta) = \kappa_i(\lambda\mu_x, \beta)$ for any $\lambda > 1, \mu_x > 0$. That is

$\kappa_i(\mu_x, \beta)$ does not depends on μ_x. Thus it could be written as $\tilde{\kappa}_i(\beta) := \kappa_i(\mu_x, \beta)$ which is a constant given a value of β, that is $\tilde{\kappa}_i(\beta) = \varsigma_i \geq 0$. The definitions of $\psi_i(\mu_x)$ and $\Psi(\mu_x)$ impose additional constraints on ς_i, namely $\varsigma_i \geq \varsigma_{i-1}$ and $\sum_{i=1}^{n} \varsigma_i = \beta$.

Conversely if $\Psi(\mu_x) = 1$, then from (102) we obtain $\psi_i(\mu_x) - \psi_i(\lambda\mu_x) = 0$ for all $\lambda > 1$, that is $\psi_i(\mu_x)$ is constant, i.e. $\psi_i(\mu_x) = \tilde{\upsilon}_i$ for all $\mu_x > 0$.

Denoting with $\tilde{\upsilon}_i := \frac{\varsigma_i}{\beta} \geq 0$ and substituting for $\psi_i(\mu_x)$ and $\Psi(\mu_x)$ into (87) we obtain:

$$\text{either } d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\frac{n\tilde{\upsilon}_i\beta + x_i}{\mu_x + \beta}\right] \quad \text{or } d_i(\mathbf{x},\varepsilon) = \varepsilon\tilde{\upsilon}_i \quad \forall i \in \mathcal{N}, \mathbf{x} \in \mathcal{X}^n. \tag{111}$$

Letting $n\tilde{\upsilon}_i = \upsilon_i$ we obtain the representation in (97) and (98). ■

5.1.11 Proposition 11

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x}, \varepsilon)$ satisfies LC, ADD, PI and WCIC if and only if for all $\mathbf{x} \in \mathcal{X}^n$, $\varepsilon > 0$: either

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\frac{\mathbf{x}}{\mu_x}, \tag{112}$$

or, there exists $\boldsymbol{\upsilon} \in \mathbb{R}_+^n$, such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$, $\sum_{i=1}^{n} \upsilon_i = n$, for any $i = 1, 2, ...n$, and

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\boldsymbol{\upsilon}. \tag{113}$$

Proof of Proposition 11:

Necessity: Consider Proposition 10. Adding WCIC implies that $\alpha d_i(\mathbf{x},\varepsilon) = d_i(\alpha\mathbf{x},\alpha\varepsilon)$ for all $\mathbf{x} \in \mathcal{X}^n$, $\varepsilon > 0$, and all $i \in \mathcal{N}$. That is from (97) and (98) we get

$$\alpha\frac{\varepsilon}{n}\left[\frac{\upsilon_i\beta + x_i}{\beta + \mu_x}\right] = \frac{\alpha\varepsilon}{n}\left[\frac{\upsilon_i\beta + \alpha x_i}{\beta + \alpha\mu_x}\right], \quad \text{and } \alpha\frac{\varepsilon}{n}\upsilon_i = \frac{\alpha\varepsilon}{n}\upsilon_i \quad \forall i \in \mathcal{N} \tag{114}$$

and for all $x_i \geq 0, \mu_x, \varepsilon, \alpha > 0$. Note that $d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\upsilon_i$ satisfies WCIC, therefore it is a solution. While the remaining solution requires to identify restrictions on the vector $\boldsymbol{\upsilon}$ and the constant $\beta \geq 0$ satisfying

$$\frac{\upsilon_i\beta + x_i}{\beta + \mu_x} = \frac{\upsilon_i\beta + \alpha x_i}{\beta + \alpha\mu_x} \quad \text{for all } x_i \geq 0,\ \mu_x, \alpha > 0,\ \text{all } i \in \mathcal{N}. \tag{115}$$

After simplifying we get

$$\alpha\mu_x\upsilon_i\beta + \beta x_i = \alpha\beta x_i + \mu_x\upsilon_i\beta \quad \text{for all } x_i \geq 0,\ \mu_x, \alpha > 0,\ \text{all } i \in \mathcal{N}. \tag{116}$$

A solution is obtained for $\beta = 0$. If $\beta \neq 0$ then

$$\alpha\upsilon_i + s_i = \alpha s_i + \upsilon_i \quad \text{for all } x_i \geq 0,\ \mu_x, \alpha > 0,\ \text{all } i \in \mathcal{N}, \tag{117}$$

where $s_i = x_i/\mu_x \geq 0$. There is no vector $\boldsymbol{\upsilon} \in \mathbb{R}_+^n$ such that the functional equation is satisfied. Therefore the general solution of (114) is obtained for $\beta = 0$.

Sufficiency: Check that (112) and (113) satisfy properties LC, ADD, PI and WCIC. ■

In order to show that LC, ADD, PI and WCIC are independent we show that if one of the axioms is dropped then a larger set of solutions is obtained. In Proposition 10 we have the characterization obtained dropping WCIC, where the cases where $\beta \neq 0$ are considered. If ADD is dropped then we obtain a class of solutions that includes those in Proposition 8. Clearly the solutions in Proposition 8 where $\lambda \neq 0, 1$, are different from those obtained in Proposition 11. If LC is dropped then $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\mathbf{w}$ is a possible solution, where $\mathbf{w} \in \mathbb{R}^n_+$ is a non-ordered vector, such that $\sum_{i=1}^n w_i = n$. This solution is not considered in Proposition 11. If finally PI is dropped then, as shown in Proposition 12, we obtain $\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\pi\frac{\mathbf{x}}{\mu_x} + (1-\pi)\boldsymbol{\upsilon}\right]$ where $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$, is an ordered vector such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$, $\sum_{i=1}^n \upsilon_i = n$. For $\pi \neq 0, 1$ these solutions differ from those obtained in Proposition 11.

5.1.12 Proposition 12

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC, ADD and WCIC if and only if there exist a ordered vector $\boldsymbol{\upsilon} \in \mathbb{R}^n_+$, s.t. $\upsilon_i \geq \upsilon_{i-1} \geq 0$ for all $i \in \mathcal{N}$, $\sum_{i=1}^n \upsilon_i = n$, and a constant $\pi \in [0, 1]$ s.t. for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$:

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\pi\frac{\mathbf{x}}{\mu_x} + (1-\pi)\boldsymbol{\upsilon}\right]. \tag{118}$$

Proof of Proposition 12:

Necessity: Consider Lemma 9. Adding WCIC requires that $\alpha d_i(\mathbf{x},\varepsilon) = d_i(\alpha\mathbf{x},\alpha\varepsilon)$ for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$, all $i \in \mathcal{N}$. From (87) in Lemma 9 we get

$$\alpha\frac{\varepsilon}{n}\left[n\psi_i(\mu_x) + (1-\Psi(\mu_x))\frac{x_i}{\mu_x}\right] = \frac{\alpha\varepsilon}{n}\left[n\psi_i(\alpha\mu_x) + (1-\Psi(\alpha\mu_x))\frac{\alpha x_i}{\alpha\mu_x}\right] \;\; \forall i \in \mathcal{N} \tag{119}$$

and for all $x_i \geq 0, \mu_x, \varepsilon, \alpha > 0$ where $\Psi(\mu_x) := \sum_{i=1}^n \psi_i(\mu_x)$. After simplifying we get

$$n\psi_i(\mu_x) + (1-\Psi(\mu_x))\frac{x_i}{\mu_x} = n\psi_i(\alpha\mu_x) + (1-\Psi(\alpha\mu_x))\frac{x_i}{\mu_x} \;\; \forall i \in \mathcal{N} \tag{120}$$

and for all $x_i \geq 0, \mu_x, \alpha > 0$. Suppose that for the individual in position $j < n$ we have $x_j = 0$ then the functional equation is satisfied only if

$$\psi_j(\mu_x) = \psi_j(\alpha\mu_x) \;\text{ for all } \mu_x, \alpha > 0.$$

If this is the case then suppose that $x_j > 0$ it follows that (120) is satisfied only if

$$\Psi(\mu_x) = \Psi(\alpha\mu_x) \quad \text{for all } \mu_x, \alpha > 0. \tag{121}$$

Now suppose this is the case it follows that (120) is satisfied only if

$$\psi_i(\mu_x) = \psi_i(\alpha\mu_x) \;\text{ for all } i \in \mathcal{N} \text{ all } \mu_x, \alpha > 0. \tag{122}$$

that implies (121). The functional equation in (122) is the necessary and sufficient condition for the solution of (120).

The general solution of (122) is obtained for $\psi_i(\mu_x) = c_i$ for all $\mu_x > 0$, where $c_i \geq c_{i-1} \geq 0$ and $1 \geq \sum_{i=1}^n c_i \geq 0$. We obtain

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[nc_i + \left(1 - \sum\nolimits_{i=1}^n c_i\right)\frac{x_i}{\mu_x}\right] \quad \forall i \in \mathcal{N}. \tag{123}$$

Letting $nc_i = \hat{\upsilon}_i$, then $\sum_{i=1}^n c_i = \sum_{i=1}^n \hat{\upsilon}_i/n$. We now consider a vector $\boldsymbol{\upsilon} \in \mathbb{R}_+^n$, such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$, and $\sum_{i=1}^n \upsilon_i = n$. For any vector $\hat{\boldsymbol{\upsilon}}$ there always exist a vector $\boldsymbol{\upsilon}$ and a constant $\pi \in [0,1]$ where $(1-\pi)\sum_{i=1}^n \upsilon_i = \sum_{i=1}^n \hat{\upsilon}_i$ i.e. $1 - \pi = \sum_{i=1}^n \hat{\upsilon}_i/n$, such that $\hat{\boldsymbol{\upsilon}} = (1-\pi)\,\boldsymbol{\upsilon}$. It follows that $\hat{\upsilon}_i = (1-\pi)\,\upsilon_i$. Substituting into (123) gives

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[\pi\frac{x_i}{\mu_x} + (1-\pi)\,\upsilon_i\right] \quad \forall i \in \mathcal{N}. \tag{124}$$

Sufficiency: Check that (118) satisfies properties LC, ADD and WCIC. ■

5.1.13 Proposition 13

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC, HE, ADD and B if and only if there exist two constants $\tilde{\omega} \in [0,1]$, $\tilde{\theta} \in \mathbb{R}_+$ such that for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$:

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left(\tilde{\omega}\frac{\mathbf{x} - \mu_x\mathbf{1}}{\mu_x + \tilde{\theta}} + \mathbf{1}\right). \tag{125}$$

Proof of Proposition 13:

Sufficiency: Check that (125) satisfies conditions LC, HE, ADD, B. Note that also VE is satisfied.

Necessity: Since for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ ADD is equivalent to $\gamma(\mu,\varepsilon) = \tilde{\gamma}(\mu)$ for all $\mu, \varepsilon > 0$, then $\lim_{\varepsilon \to 0^+} \gamma(\mu,\varepsilon)$ exists and is $\tilde{\gamma}(\mu)$. In order to be consistent with the notation introduced in Proposition 2 we let $\eta(\mu) := \tilde{\gamma}(\mu)$. From Lemma 7 we know that if $\eta(\mu)$ exists then, for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfying B, it corresponds to $\eta(\mu) = \tilde{\omega}\frac{\mu}{\mu+\tilde{\theta}}$ where $\tilde{\theta} \geq 0$, $\tilde{\omega} \in [0,1]$. Therefore for all $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_{\mathcal{E}}(\mathbf{x},\varepsilon)$ satisfying ADD+B we have $\gamma(\mu,\varepsilon) = \tilde{\omega}\frac{\mu}{\mu+\tilde{\theta}}$ for all $\mu, \varepsilon > 0$. From Remark 4 property VE can be dropped since it is implied by LC+ADD. ■

5.1.14 Proposition 14

First we prove the following lemma:

Lemma 10 *An IEDV* $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ *satisfies conditions LC and WCIC for all* $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$, *if and only if there exist a set of continuous functions* $\phi_i : \mathbb{R}_{++} \to \mathbb{R}_+$ *such that (i)* $\phi_j(\mu/\varepsilon) \geq \phi_i(\mu/\varepsilon) \geq 0$ *if* $x_j \geq x_i$ *for all* $\mu,\varepsilon > 0$, $i,j \in \mathcal{N}$ *, (ii)* $1 - \sum_{h=1}^n \phi_h(\mu/\varepsilon) \geq -\min_i\{\phi_i(\mu/\varepsilon)\cdot(n-i+1) : i \in \mathcal{N}\}$ *where* $x_{i+1} \geq x_i$, *for all* $\mu, \varepsilon > 0$, *(iii)* $1 - \sum_{h=1}^n \phi_h(\mu/\varepsilon) \geq -n(\mu/\varepsilon)$ *for all* $\mu, \varepsilon > 0$ *and*

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[n\phi_i(\mu_x/\varepsilon) + [1 - \Phi(\mu_x/\varepsilon)]\frac{x_i}{\mu_x}\right] \quad \forall i \in \mathcal{N}$$

where $\Phi(\mu/\varepsilon) := \sum_{i=1}^n \phi_i(\mu/\varepsilon)$.

Proof. *Necessity:* Letting $a_i(\mu,\varepsilon) = g_i(\mu,\varepsilon)\varepsilon$ in Lemma 1, we derive the IEDVs $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfying LC:

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[ng_i(\mu,\varepsilon) + \frac{1-\sum_{j=1}^n g_j(\mu,\varepsilon)}{\mu}x_i\right]$$

where (i) $g_j(\mu,\varepsilon) \geq g_i(\mu,\varepsilon) \geq 0$ if $x_j > x_i$ for all $\mu,\varepsilon > 0$, (ii) $1+\min_i \{g_i(\mu,\varepsilon)\cdot(n-i+1) : i \in \mathcal{N}\} \geq \sum_{h=1}^n g_h(\mu,\varepsilon)$ for all $\mu,\varepsilon > 0$, where $x_{i+1} \geq x_i$, and (iii) $1 + n\mu/\varepsilon \geq \sum_{h=1}^n g_h(\mu,\varepsilon)$, for all $\mu,\varepsilon > 0$.

Applying WCIC we get

$$\begin{aligned} \alpha\frac{\varepsilon}{n}\left[ng_i(\mu,\varepsilon) + \left[1-\sum\nolimits_{j=1}^n g_i(\mu,\varepsilon)\right]\frac{x_i}{\mu}\right] &= \alpha d_i(\mathbf{x},\varepsilon) = \\ \frac{\alpha\varepsilon}{n}\left[ng_i(\alpha\mu,\alpha\varepsilon) + \left[1-\sum\nolimits_{j=1}^n g_i(\alpha\mu,\alpha\varepsilon)\right]\frac{\alpha x_i}{\alpha\mu}\right] &= d_i(\alpha\mathbf{x},\alpha\varepsilon) \end{aligned}$$

that is

$$ng_i(\mu,\varepsilon) + \left[1-\sum\nolimits_{j=1}^n g_i(\mu,\varepsilon)\right]\frac{x_i}{\mu} = ng_i(\alpha\mu,\alpha\varepsilon) + \left[1-\sum\nolimits_{j=1}^n g_i(\alpha\mu,\alpha\varepsilon)\right]\frac{x_i}{\mu}$$

for all $\mathbf{x} \in X^n$, all $\alpha,\varepsilon > 0$.

Considering the two distributions $\mathbf{x}' = (0,0,0,0,...,0,n\mu)$, and $\mathbf{x}'' = \mu\mathbf{1}$ it is possible to show that the previous functional equations boils down to:

$$g_i(\mu,\varepsilon) = g_i(\alpha\mu,\alpha\varepsilon)$$

for all $\alpha,\varepsilon,\mu > 0$. Whose solution is obtained letting $\alpha = 1/\varepsilon$ and defining $\phi_i(\mu/\varepsilon) := g_i(\mu/\varepsilon,1)$ for all $\varepsilon,\mu > 0$.

Note that for all $t > 0$ the previous conditions become (i) $\phi_j(t) \geq \phi_i(t) \geq 0$ if $x_j > x_i$, (ii) $1 - \sum_{h=1}^n \phi_h(t) \geq -\min_i \{\phi_i(t)\cdot(n-i+1) : i \in \mathcal{N}\}$ where $x_{i+1} \geq x_i$, and (iii) $1 - \sum_{h=1}^n \phi_h(t) \geq -nt$.

Sufficiency: Check that properties LC and WCIC are satisfied. ■

Adding HE we get:

(i) $\phi_i(\mu/\varepsilon) = \phi(\mu/\varepsilon) \geq 0$ for all $i \in \mathcal{N}$, (ii) $\frac{1}{n-1} \geq \phi(\mu/\varepsilon)$, for all $\mu,\varepsilon > 0$, (iii) $\frac{1}{n} + (\mu/\varepsilon) \geq \phi(\mu/\varepsilon)$ for all $\mu,\varepsilon > 0$ and

$$d_i(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\left[n\phi(\mu_x/\varepsilon) + [1-n\phi(\mu_x/\varepsilon)]\frac{x_i}{\mu_x}\right] \quad \forall i \in \mathcal{N}.$$

Proposition 14: An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies conditions LC, PI and WCIC if and only if there exist $\boldsymbol{\upsilon} \in \mathbb{R}_+^n$ such that $\upsilon_i \geq \upsilon_{i-1} \geq 0$ for all $i \in \mathcal{N}$, $\sum_{i=1}^n \upsilon_i = n$, and a constant $\lambda \in [0,1]$ such that for all $\mathbf{x} \in \mathcal{X}^n, \varepsilon > 0$:

$$\mathbf{d}(\mathbf{x},\varepsilon) = \frac{\varepsilon}{n}\boldsymbol{\upsilon} + (\mathbf{x} - \mu_x\boldsymbol{\upsilon})\left[\left(1 + \frac{\varepsilon}{n}\cdot\frac{1}{\mu_x}\right)^\lambda - 1\right]. \tag{126}$$

Proof of Proposition 14:

Necessity: In order to simplify the calculations, let $h_i(t) := 1 - n\phi_i(1/t)$ for all $t > 0$ i.e. $\phi_i(t) = \frac{1}{n}[1 - h_i(1/t)]$. Applying Lemma 3 if $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}(\mathbf{x},\varepsilon)$ satisfies LC and PI then $\lim_{\varepsilon\to 0^+} h_i(\varepsilon/\mu)$ exists implying that $\lim_{\varepsilon\to 0^+} \frac{d_i(\mathbf{x},\varepsilon)}{\varepsilon} = \delta_i(\mathbf{x}) \geq 0$ exists and is:

$$\begin{aligned}
\delta_i(\mathbf{x}) &= \lim_{\varepsilon\to 0^+} \frac{1}{n}\left[[1 - h_i(\varepsilon/\mu)] + \frac{1}{n}\left[\sum\nolimits_{j=1}^n h_j(\varepsilon/\mu)\right]\frac{x_i}{\mu_x}\right] \\
&= \frac{1}{n}\left[\left[1 - \lim_{\varepsilon\to 0^+} h_i(\varepsilon/\mu)\right] + \frac{1}{n}\left[\sum\nolimits_{j=1}^n \lim_{\varepsilon\to 0^+} h_j(\varepsilon/\mu)\right]\frac{x_i}{\mu_x}\right] \\
&= \frac{1}{n}\left[[1 - \bar{\eta}_i] + \frac{1}{n}\left[\sum\nolimits_{j=1}^n \bar{\eta}_j\right]\frac{x_i}{\mu_x}\right]
\end{aligned}$$

where $\bar{\eta}_i := \lim_{\varepsilon\to 0^+} h_i(\varepsilon/\mu)$ for all $\mu > 0$. We get

$$n\delta_i(\mathbf{x}) = [1 - \bar{\eta}_i] + \left[\sum\nolimits_{j=1}^n \bar{\eta}_i\right]\frac{x_i}{n\mu}$$

where $1 \geq \bar{\eta}_i \geq \bar{\eta}_{i+1}$ and $\sum_{j=1}^n \bar{\eta}_i \geq 0$. Let $\sum_{j=1}^n \bar{\eta}_i = \lambda n \geq 0$, where $\lambda \geq 0$, and denote $\upsilon_i := \frac{1-\bar{\eta}_i}{(1-\lambda)}$, notice that $1 - \bar{\eta}_i = \upsilon_i(1-\lambda)$, that is $\sum_{i=1}^n (1-\bar{\eta}_i) = \sum_{i=1}^n \upsilon_i(1-\lambda)n$, giving $n - n\lambda = (1-\lambda)\sum_{i=1}^n \upsilon_i$, it follows that $\sum_{i=1}^n \upsilon_i = n$, and from $1 \geq \bar{\eta}_i \geq \bar{\eta}_{i+1}$ follows that $\upsilon_i \geq \upsilon_{i-1} \geq 0$. Substituting we get

$$n\delta_i(\mathbf{x}) = \upsilon_i(1-\lambda) + [\lambda n]\frac{x_i}{n\mu} = \upsilon_i(1-\lambda) + \lambda\frac{x_i}{\mu}.$$

Recalling from the proof of Proposition 2 that $n \cdot \boldsymbol{\delta}(\mathbf{x}) = \dot{\mathbf{x}}$ then we obtain

$$\dot{\mathbf{x}} = \boldsymbol{\upsilon}(1-\lambda) + \lambda\frac{\mathbf{x}}{\mu_x} \tag{127}$$

where $\boldsymbol{\upsilon} \in \mathbb{R}_+^n$ is the vector of components $\upsilon_i \geq \upsilon_{i-1} \geq 0$, such that $\sum_{i=1}^n \upsilon_i = n$. Equivalently we have

$$\dot{x}_i(\mu) = \left[\upsilon_i + \left(\frac{x_i(\mu) - \upsilon_i\mu}{\mu}\right)\lambda\right] \quad \forall i \in \mathcal{N}. \tag{128}$$

For the solution of the system of differential equations we follow the same logic as applied in the proof of Proposition 2. Denote $z_i(\mu) = x_i(\mu) - \upsilon_i\mu$; noticing that $\dot{z}_i = \dot{x}_i - \upsilon_i$ we can rewrite (128) as

$$\dot{z}_i(\mu) = \lambda\frac{z_i(\mu)}{\mu} \tag{129}$$

multiply both sides of this equation by $\exp(-\int_{\mu_0}^{\mu}\frac{\lambda}{\tau}d\tau)$, and let $\exp(-\int_{\mu_0}^{\mu}\frac{\lambda}{\tau}d\tau)\cdot z_i(\mu) = h_i(\mu)$. It follows that :

$$\dot{h}_i(\mu) = -\frac{\lambda}{\mu}\exp\left(-\int_{\mu_0}^{\mu}\frac{\lambda}{\tau}d\tau\right)\cdot z_i(\mu) + \exp\left(-\int_{\mu_0}^{\mu}\frac{\lambda}{\tau}d\tau\right)\cdot\dot{z}_i(\mu). \tag{130}$$

Since $\dot{z}_i = \frac{\lambda}{\mu} z_i$, then $\dot{h}_i(\mu) = 0$, thus

$$\exp\left(-\int_{\mu_0}^{\mu} \frac{\lambda}{\tau} d\tau\right) \cdot z_i(\mu) = K \tag{131}$$

where K is an arbitrary constant. That is, in terms of the distances of x_i from the weighted mean income $\upsilon_i \mu_x$ we have:

$$x_i - \upsilon_i \mu_x = K \cdot \exp\left(\int_{\mu_0}^{\mu_x} \frac{\lambda}{\tau} d\tau\right) \quad \text{for all } i \in \mathcal{N}. \tag{132}$$

It follows that

$$x_i - \upsilon_i \mu_x = K \cdot \exp\left\{\left[\ln\left(\mu^{\lambda}\right)\right]_{\mu_0}^{\mu_x}\right\} \quad \text{for all } i \in \mathcal{N} \tag{133}$$

that is

$$x_i - \upsilon_i \mu_x = K \cdot \left[\frac{\mu_x}{\mu_0}\right]^{\lambda} = (\mu_x)^{\lambda} \cdot \frac{K}{(\mu_0)^{\lambda}} \quad \text{for all } i \in \mathcal{N}. \tag{134}$$

Thus, for all distributions $\mathbf{x} \in \mathcal{X}^n$

$$\frac{x_i - \upsilon_i \mu_x}{(\mu_x)^{\lambda}} = \bar{K} = \frac{K}{(\mu_0)^{\lambda}} \quad \text{for all } i \in \mathcal{N}, \tag{135}$$

that is

$$\frac{x_i - \upsilon_i \mu_x}{(\mu_x)^{\lambda}} = \frac{y_i - \upsilon_i \mu_y}{(\mu_y)^{\lambda}} \quad \text{for all } i \in \mathcal{N}, \text{ all } \mathbf{x}, \mathbf{y} \in \mathcal{X}^n \tag{136}$$

where $\lambda \geq 0$. Rearranging supposing that $\mu_y > \mu_x$ i.e. $\mu_y = \mu_x + \frac{\varepsilon}{n}$ we get

$$\begin{aligned} d_i(\mathbf{x}, \varepsilon) &= (y_i - x_i) = \upsilon_i \mu_y + \left(\frac{\mu_y}{\mu_x}\right)^{\lambda} (x_i - \upsilon_i \mu_x) - x_i \\ &= \upsilon_i \left(\mu_x + \frac{\varepsilon}{n}\right) + \left(\frac{\mu_x + \frac{\varepsilon}{n}}{\mu_x}\right)^{\lambda} (x_i - \upsilon_i \mu_x) - x_i \\ &= \upsilon_i \frac{\varepsilon}{n} + \left[\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda} - 1\right] (x_i - \upsilon_i \mu_x), \end{aligned}$$

for all $i \in \mathcal{N}$, that is

$$\mathbf{d}(\mathbf{x}, \varepsilon) = \frac{\varepsilon}{n} \boldsymbol{\upsilon} + (\mathbf{x} - \mu_x \boldsymbol{\upsilon}) \left[\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda} - 1\right] \quad \text{where } \lambda \geq 0. \tag{137}$$

Note that $\mathbf{d}(\mathbf{x}, \varepsilon) \geq 0$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}^n$ requires that $\lambda \leq 1$. Suppose that $x_i = 0$, then $d_i(\mathbf{x}, \varepsilon) = \upsilon_i \left[\frac{\varepsilon}{n} - \mu_x \left[\left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda} - 1\right]\right]$ where $\upsilon_i \geq 0$, it follows that $d_i(\mathbf{x}, \varepsilon) \geq 0$ implies that $1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x} \geq \left(1 + \frac{\varepsilon}{n} \cdot \frac{1}{\mu_x}\right)^{\lambda}$ that is $\lambda \leq 1$.

Sufficiency: Check that (126) satisfies properties LC, PI and WCIC. ■

To complete we check that (126) satisfies VE i.e. $d_i(\mathbf{x},\varepsilon) \geq d_{i-1}(\mathbf{x},\varepsilon)$ for all $i \in \mathcal{N}$, all $\mathbf{x} \in \mathcal{X}^n$, all $\varepsilon > 0$. We check this condition when $x_i = x_{i-1}$, if VE is satisfied in this case then it will be also satisfied in all the cases where $x_i > x_{i-1}$. According to (126) if $x_i = x_{i-1}$ we get

$$\begin{aligned} d_i(\mathbf{x},\varepsilon) - d_{i-1}(\mathbf{x},\varepsilon) &= (\upsilon_i - \upsilon_{i-1})\left\{\frac{\varepsilon}{n} - \mu_x\left[\left(1+\frac{\varepsilon}{n}\cdot\frac{1}{\mu_x}\right)^{\lambda} - 1\right]\right\} \\ &= (\upsilon_i - \upsilon_{i-1})\cdot\mu_x\cdot\left[\left(1+\frac{\varepsilon}{n}\cdot\frac{1}{\mu_x}\right) - \left(1+\frac{\varepsilon}{n}\cdot\frac{1}{\mu_x}\right)^{\lambda}\right] \end{aligned}$$

where the last term on the r.h.s. is non-negative if $\lambda \leq 1$. It follows that $d_i(\mathbf{x},\varepsilon) - d_{i-1}(\mathbf{x},\varepsilon) \geq 0$, that is VE is implied by the properties applied.

5.1.15 Proposition 15

An IEDV $\mathbf{d}(\mathbf{x},\varepsilon) \in \mathcal{D}_R(\mathbf{x},\varepsilon)$ satisfies LC if and only if there exists a sequence of continuous functions $a_i(\mu,\varepsilon) : \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}_+$, satisfying $\varepsilon = \sum_{i=1}^{n} a_i(\mu,\varepsilon)$, such that for all $\mathbf{x} \in \mathbb{R}^n$ and $\varepsilon > 0$:

$$d_i(\mathbf{x},\varepsilon) = a_i(\mu,\varepsilon) \qquad \forall i = 1,2,..n$$

where $a_i(\mu,\varepsilon) \geq a_j(\mu,\varepsilon)$ if $x_i > x_j$.

Proof. The sufficiency party is checked noticing that all distributions in $\mathcal{P}(\mathbf{x})$ have the same average thus when the same surplus is distributed we have $d_i(\mathbf{x},\varepsilon) = a_i(\mu,\varepsilon) = d_i(\mathbf{y},\varepsilon)$ for all $\mathbf{y} \in \mathcal{P}(\mathbf{x})$. It follows that $\mathbf{y}' - \mathbf{x}' = \mathbf{y} - \mathbf{x}$ where $\mathbf{y}' \in \mathcal{P}(\mathbf{x}')$, and $x_i + d_i(\mathbf{x},\varepsilon) \geq x_j + d_j(\mathbf{x},\varepsilon)$ if $x_i \geq x_j$. Therefore LC is satisfied. Moreover the vector of elements $a_i(\mu,\varepsilon)$ belongs to $\mathcal{D}_R(\mathbf{x},\varepsilon)$ since they are continuous, non negative and satisfy $\varepsilon = \sum_{i=1}^{n} a_i(\mu,\varepsilon)$.

The proof of the necessity part follows from Lemma 1, here we have substituted $\mathbf{x} \in \mathbb{R}^n$ for $\mathbf{x} \in X^n$. Some results derived there still hold for the more general case where $\mathbf{x} \in \mathbb{R}^n$. As in the proof of Lemma 1, applying LC to the IEDVs in $\mathcal{D}_R(\mathbf{x},\varepsilon)$ we derive the restriction

$$d_h(\mathbf{x},\varepsilon) = f_h(x_h,\mu,\varepsilon),$$

where $x_h \in \mathbb{R}$ is the income of agent h, $\mu \in \mathbb{R}$ is average income of distribution $\mathbf{x} \in \mathbb{R}^n$, and the function $f_h(.)$ is continuous, and satisfies $\sum_h f_h(.) = \varepsilon$. As discussed in the proof of Lemma 1, given the definition of LC, for any rank-preserving transfer $\delta \geq 0$ we have $f_i(x_i+\delta,\mu,\varepsilon) \geq f_i(x_i,\mu,\varepsilon) - \delta$, and $f_j(x_j-\delta,\mu,\varepsilon) \leq f_j(x_j,\mu,\varepsilon) + \delta$ and if $x_i > x_j$ then $x_i + d_i(\mathbf{x},\varepsilon) > x_j + d_j(\mathbf{x},\varepsilon)$. Recalling that $\sum_h f_h(.) = \varepsilon$ we obtain

$$f_i(x_i+\delta,\mu,\varepsilon) + f_j(x_j-\delta,\mu,\varepsilon) = f_j(x_j,\mu,\varepsilon) + f_i(x_i,\mu,\varepsilon)$$

for any i,j; for all $x_i, x_j \in \mathbb{R}$ and any rank-preserving transfer of the amount $\delta \geq 0$. Letting $g_i(x_i,\delta,\mu,\varepsilon) = f_i(x_i+\delta,\mu,\varepsilon) - f_i(x_i,\mu,\varepsilon)$, then

$$g_i(x_i,\delta,\mu,\varepsilon) = g_j(x_j-\delta,\delta,\mu,\varepsilon)$$

for any $i, j; x_i < x_j$ and any rank-preserving transfer of amount $\delta \geq 0$. Following the same line of reasoning as in the proof of Lemma 1 it turns out that $g_i(x_i, \delta, \mu, \varepsilon) = g(\delta, \mu, \varepsilon)$. Thus,

$$f_i(x_i, \mu, \varepsilon) + g(\delta, \mu, \varepsilon) = f_i(x_i + \delta, \mu, \varepsilon) \tag{138}$$

for any $\delta \geq 0$, where $g(0, \mu, \varepsilon) = 0$, but differently from what supposed in Lemma 1 here we have $x \in \mathbb{R}$. The functional equation associated with (138) can be rewritten as

$$F(x) + G(\delta) = F(x + \delta) \quad \text{for all } \delta \in \mathbb{R}_+, \text{ all } x \in \mathbb{R},$$

where $F(x) := f_i(x_i, \mu, \varepsilon)$, and $G(\delta) := g(\delta, \mu, \varepsilon)$. Letting $x = 0$, then

$$\alpha + G(\delta) = F(\delta), \quad \text{for all } \delta \in \mathbb{R}_+, \text{ where } \alpha = F(0),$$

therefore $G(\delta) = F(\delta) - \alpha$ for all $\delta \in \mathbb{R}_+$, leading to

$$F(x) + F_+(\delta) - \alpha = F(x + \delta)$$

where $F_+(.)$ is the restriction of $F(.)$ defined over the domain of non negative real numbers. Letting $H(x) = F(x) - \alpha$, then we can write

$$H(x) + H_+(\delta) = H(x + \delta) \quad \text{for all } \delta \in \mathbb{R}_+, \text{ all } x \in \mathbb{R}. \tag{139}$$

This is a general version of the Cauchy functional equation, since $H_+(.)$ is defined over a more restricted domain compared to $H(.)$. Eichhorn (1978) (Th. 2.6.3 p.39) provides a solution for the general version of Cauchy functional equation considering functions defined over possibly different arbitrary non degenerate intervals of $\mathbb{R}$. The result for the case of functions defined on two intervals is:

Theorem 1 (Eichhorn, 1978 Th. 2.6.3) *The general solution of the functional equation $\phi(x + y) = \psi_1(x) + \psi_2(y)$ for all $x \in I_1, y \in I_2$, and $(x + y) \in I_1 + I_2$ where I_1, I_2 are arbitrary non degenerate intervals of $\mathbb{R}$ and the functions $\phi(.), \psi_1(.)$ and $\psi_2(.)$ are continuous is given by*

$$\begin{aligned} \phi(x + y) &= \beta(x + y) + \gamma_1 + \gamma_2 \\ \psi_1(x) &= \beta x + \gamma_1, \qquad \psi_2(y) = \beta y + \gamma_2 \end{aligned}$$

for $\beta, \gamma_1, \gamma_2$ real constants.

Proof of Proposition 15 (Continued): In our case $\psi_2(y) = H(y) = \phi(y)$ therefore $\gamma_1 = 0$, moreover $\psi_1(x) = H_+(x) = \phi(x)$ whenever $x \in \mathbb{R}_+$, from which follows that also $\gamma_2 = 0$. As a result the general solution of (139) is the same as that for the Cauchy equation defined over real numbers or non negative real numbers. Substituting for

$$H(y) = \beta y, \ H_+(x) = \beta x, \quad \beta \in \mathbb{R}, y \in \mathbb{R}, x \in \mathbb{R}_+$$

we get

$$F(y) = \beta y + \alpha, \quad F_+(x) - \alpha = G(x) = \beta x, \quad \beta, \alpha \in \mathbb{R}, y \in \mathbb{R}, x \in \mathbb{R}_+.$$

That is, since $G(\delta) := g(\delta, \mu,\varepsilon)$ and $F(x) := f_i(x_i, \mu,\varepsilon)$, we have that $\beta := b(\mu,\varepsilon)$ while $\alpha := a_i(\mu,\varepsilon)$ getting

$$\begin{aligned} G(x) &= b(\mu,\varepsilon)x, \ F(x) = b(\mu,\varepsilon)x + a_i(\mu,\varepsilon), \\ &\quad \text{where } \ b(\mu,\varepsilon) \text{ and } a_i(\mu,\varepsilon) \text{ are continuous real functions.} \end{aligned}$$

The general solution of (138) is therefore the same as obtained in Lemma 1:

$$f_i(x_i, \mu,\varepsilon) = a_i(\mu,\varepsilon) + b(\mu,\varepsilon)x_i \quad \text{for all } \mu,\varepsilon > 0, \ x_i \in \mathbb{R}. \tag{140}$$

Consider now condition $f_h(.) \geq 0$. Setting $x_i = 0$, it follows that $a_i(\mu,\varepsilon) \geq 0$. Moreover, it must be $b(\mu,\varepsilon) = 0$, since otherwise, for every $a_i(\mu,\varepsilon) \geq 0$ there will always exists a value x_i^* of x_i such that $f_h(.) < 0$. This is the case if $x_i^* < -a_i(\mu,\varepsilon)/b(\mu,\varepsilon)$ when $b(\mu,\varepsilon) > 0$, and $x_i^* > -a_i(\mu,\varepsilon)/b(\mu,\varepsilon)$ when $b(\mu,\varepsilon) < 0$.

Considering the requirement $\sum_h f_h(.) = \varepsilon$ we get $\sum_{i=1}^n a_i(\mu,\varepsilon) = \varepsilon$, thus:

$$d_i(\mathbf{x},\varepsilon) = a_i(\mu,\varepsilon) \ \text{ where } \sum\nolimits_{i=1}^{n} a_i(\mu,\varepsilon) = \varepsilon \ , \tag{141}$$

for all $i = 1, 2, ..n$, and all $\mu \in \mathbb{R}$, $\varepsilon > 0$.

Finally the requirement that if $x_i > x_j$ then $x_i + d_i(\mathbf{x}, \varepsilon) > x_j + d_j(\mathbf{x}, \varepsilon)$ implies that $a_i(\mu,\varepsilon) \geq a_j(\mu,\varepsilon)$ if $x_i \geq x_j$. ■