Society for the Study of Economic Inequality

Working Paper Series

Inequality measurement for bounded variables

Inaki Permanyer

Suman Seth

Gaston Yalonetzky

ECINEQ 2022 602

ECINEQ 2022 602
2022 February
www.ecineq.org

Inequality measurement for bounded variables

Inaki Permanyer
ICREA Centre d'Estudis Demogràfics
Suman Seth
University of Leeds
Gaston Yalonetzky
University of Leeds

Abstract

Numerous non-pecuniary variables of interest for inequality assessment are bounded and often represented in terms of attainments or shortfalls. Inequality measurement for bounded variables suffers from two key challenges: the consistency problem and the boundary problem. The former occurs when inequality rankings reverse while switching between attainment and shortfall representations. The latter stems from the existence of a predictable functional relationship between mean attainment and maximum feasible inequality hindering inequality comparisons across distributions with different means. Unlike consistency, the boundary problem has not received significant attention in the literature. We propose two novel classes of normalized inequality measures that are immune to both problems. We illustrate the empirical relevance ofour approach with cross-country comparisons of inequality in well-established indicators of education and health. A starkly different picture emerges when traditional inequality indices give way to our normalized inequality indices.

Keyword: Inequality measurement, bounded variables, boundary problem, consistency, Kuznets curves.

JEL Cassification: D63, I31, O57.

Inequality measurement for bounded variables[*]

Iñaki Permanyer[†] Suman Seth[‡] Gaston Yalonetzky[§]

February 10, 2022

Abstract

Numerous non-pecuniary variables of interest for inequality assessment are bounded and often represented in terms of attainments or shortfalls. Inequality measurement for bounded variables suffers from two key challenges: the *consistency problem* and the *boundary problem*. The former occurs when inequality rankings reverse while switching between attainment and shortfall representations. The latter stems from the existence of a predictable functional relationship between mean attainment and maximum feasible inequality hindering inequality comparisons across distributions with different means. Unlike consistency, the boundary problem has not received significant attention in the literature. We propose two novel classes of normalized inequality measures that are immune to both problems. We illustrate the empirical relevance of our approach with cross-country comparisons of inequality in well-established indicators of education and health. A starkly different picture emerges when traditional inequality indices give way to our normalized inequality indices.

Keywords: Inequality measurement, bounded variables, boundary problem, consistency, Kuznets curves.

JEL Codes: D63, I31, O57.

[*]We would like to thank Antonio Villar, Juan Moreno-Ternero, Florent Bresson, Ohiana Aristondo, Casilda Lasso de la Vega, Gordon Anderson, Jacques Silber, Stephan Klasen, Shatakshee Dhondge, Conchita D'Ambrosio and Sanghamitra Bandyopadhyay for very helpful comments and suggestions. We are also grateful to the participants of the 16th Annual Conference on Economic Growth and Development at the Indian Statistical Institute, Delhi; Economics Seminar at Universidad Pablo de Olavide; 36th General Conference of the International Association for Research in Income and Wealth; 14th Meeting of the Society for Social Choice and Welfare, Seoul, South Korea; 13th conference of the Western Economic Association International, Catholic University, Santiago, Chile; the 13th Meeting of the Society for Social Choice and Welfare, Lund University, Sweden; the 7th Meeting of the Society for the Study of Economic Inequality, CUNY, USA; and the Royal Economic Society conference, University of Sussex, UK, for their valuable comments. Permanyer acknowledges funding from the European Research Council (ERC-2019-CoG-864616, HEALIN project), the CERCA Programme from Generalitat de Catalunya and from the Spanish Ministry of Economy and Competitiveness "Ramón y Cajal" Research Grant Program (RYC-2013-14196). All remaining errors are our own.

[†]ICREA and Centre d'Estudis Demografics, Universitat Autonoma de Barcelona. Email: ipermanyer@ced.uab.es.

[‡]Economics Department, Leeds University Business School, University of Leeds and Oxford Poverty & Human Development Initiative (OPHI), University of Oxford, Oxford, UK. Email: S.Seth@leeds.ac.uk.

[§]Economics Department, Leeds University Business School, University of Leeds, University of Leeds. Email: G.Yalonetzky@leeds.ac.uk.

1 Introduction

In his seminal contribution, Atkinson (1970) set the foundations of inequality measurement as we know it. After five decades, the contributions to this burgeoning field of research have expanded in multiple directions, and 'inequality' can arguably be considered one of the most hotly debated topics in an increasingly globalised world, as witnessed by the popularity of several recent books on the subject (e.g. Piketty, 2015; Bourguignon, 2017; Atkinson, 2018; Milanovic, 2018). Moreover, the interest in inequality has gone well beyond the study of monetary distributions. Nowadays scholars and policy-makers alike are particularly interested in studying the distribution of the non-pecuniary dimensions of well-being (e.g. health and education outcomes, or the post-2015 Sustainable Development Goals agenda).

Bounded variables abound.[1] Unlike monetary variables, such as income or consumption expenditure, the majority of non-pecuniary aspects of well-being—such as literacy rates or infant mortality rates—are gauged by variables that cannot take indefinitely large values. This seemingly unimportant and technical point has key implications as regards the way in which we measure and interpret 'inequality' in the corresponding distributions. The measurement of inequality for bounded variables poses two specific challenges. While one has been properly addressed in the literature, the other has been largely neglected. We propose an approach to inequality measurement with bounded variables that addresses both challenges simultaneously.

The first problem that researchers encounter when studying inequality for bounded variables is the 'consistency problem'. When a variable is bounded, one may choose to focus either on the distribution of attainments or the corresponding distribution of shortfalls with respect to the upper bound.[2] Many inequality measures (especially popular relative measures like the Gini coefficient) fail to rank distributions consistently when measurement is switched from attainments to shortfall representations (see, among others, Micklewright and Stewart, 1999; Kenny, 2004; Clarke et al., 2002; Erreygers, 2009; Lambert and Zheng, 2011; Lasso de la Vega and Aristondo, 2012; Bosmans, 2016). Far from being a mere academic curiosity, the consistency problem poses several practical challenges to the study of inequality for bounded variables by precluding recourse to popular tools like the Lorenz curve. Fortunately, a battery of satisfactory solutions enabling the consistent measurement of inequality with bounded variables has been proposed, e.g. using absolute inequality measures (Erreygers, 2009; Lambert and Zheng, 2011), indices based on both representations (Lasso de la Vega and Aristondo, 2012), or using pairs of weakly consistent indices (Bosmans, 2016).

The second challenge is the 'boundary' problem (in some contexts also known as 'floor-' and 'ceiling-effect' problems). Whenever a variable is bounded, one can observe a clustering of the distribution as its mean converges towards any of its bounds. In these situations, the corresponding inequality levels mechanically converge toward zero, simply because there is no room left for further variation. This problem persists even when the consistency problem is solved using any of the aforementioned solutions. For instance, in the case of several absolute inequality measures and the absolute Lorenz curve, as the mean of the distribution increases from the lower bound to the upper bound of the distribution, the level of maximum feasible inequality first increases and then decreases; making the maximum feasible inequality a non-monotonic function of the mean. In these circumstances, it is not clear whether studying inequality with a bounded variable can provide new insights above and beyond what we already know from studying the values of

[1]Examples include indicators of education, health, political freedom, democracy level, freedom from violence, happiness, trust, corruption, household or environmental characteristics, access to public services, poverty, socio-demographic characteristics, etc.

[2]For instance, improvements in the coverage of public health plans could be assessed via either the percentage of vaccinated children (an achievement indicator) or the percentage of unvaccinated children (a shortfall indicator).

the mean alone.[3] More generally, for any given consistent inequality measure, maximum feasible inequality becomes a predictable function of mean attainment, which hinders the disentanglement of 'mechanical' changes from inequality dynamics due to socioeconomic phenomena when comparing distributions with different means.

We propose two new closely-related classes of inequality measures: the so-called classes of *normalized inequality measures*. They satisfy the basic requirements of inequality measurement and are not affected either by inconsistencies or the boundary problem. Both classes are obtained by combining normalization axioms with the strong consistency requirement (Bosmans, 2016) and minimum desirable properties (e.g. the transfers principle). The key distinction between these two classes is that one is defined for fixed population, while the other class is defined for variable population.

Our proposed classes are quite broad. Intuitively, their members are defined as the ratio of two identical symmetric S-convex functions such that the function in the denominator is evaluated at a distribution maximising its value and sharing the same mean as the distribution used to evaluate the function in the numerator, namely the distribution whose inequality is being measured. Hence the normalized indices lie between zero and one. The admissible functional forms for the symmetric and S-convex function include (but are not limited to) numerous absolute, relative and intermediate inequality indices. Whenever we use an inequality index as an admissible functional form, we obtain a new inequality measure quantifying the observed inequality level as a proportion of the maximum inequality level that could be attained with the same index evaluated at a hypothetical distribution with the same mean as the observed distribution.

We illustrate the empirical performance of the normalized inequality measures proposed in this paper comparing the normalized Gini index against its absolute and relative counterparts, which fail to provide inequality rankings that are simultaneously consistent and free of the boundary problem. We study the evolution of cross-country inequality in two indicators of child survival and three indicators of education-level completion since 1950. As expected, when we use the absolute Gini index we find Kuznets curves, whereby cross-country inequality initially increases with higher mean attainment and then decreases as the latter approaches its upper bound. Instead, when we use the relative Gini index, inequality steadily decreases as mean attainment increases. By stark contrast, the normalized Gini index reports increases in inequality over the same periods of high mean attainment reaching its upper limit, thereby suggesting that maximum possible Gini inequality has decreased faster than both absolute and relative Gini inequality. Moreover, the empirical application clearly illustrates that, unlike the cases of the absolute and relative Gini index, the value of their normalized version is not a predictable function of the mean.

The rest of the paper proceeds as follows. Section 2 introduces the basic inequality measurement framework including notation and basic definitions. In sections 3 and 4 we examine, respectively, the challenges of consistency and boundary problems when measuring inequality with bounded variables. Section 5 presents two classes of normalized inequality measures that provide consistent inequality comparisons free of the boundary problem. Section 6 presents the empirical illustration and section 7 concludes.

[3]This point was already highlighted by Wagstaff (2005) in his study of the concentration index, and discussed by many others after him (e.g. Erreygers, 2009; Erreygers and Van Ourti, 2011; Wagstaff, 2009, 2011; Prados de la Escosura, 2021). In this paper we extrapolate some of these ideas to the context of inequality measurement for bounded variables.

2 Basic framework for measuring inequality with bounded variables

Let the sets of real numbers, rational numbers and natural numbers be $\mathbb{R}$, $\mathbb{Q}$ and $\mathbb{N}$, respectively. The non-negative and strictly positive counterparts of $\mathbb{R}$ and $\mathbb{Q}$ are represented by adding to either the subscripts $_+$ and $_{++}$, respectively. Suppose, there are n units of analysis (e.g. people, households, municipalities, countries, etc.) such that $n \in \mathbb{N} \setminus \{1\}$. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be an *attainment distribution* of n units (or an n-dimensional *attainment vector*), where $x_i \in [0,U] \subset \mathbb{R}$ represents unit i's cardinally measurable attainment bounded between a lower bound of zero and some fixed positive upper bound $U \in \mathbb{R}_{++}$.

We denote the set of all attainment distributions of size n with upper bound U by $\mathscr{X}_n$ and the set of all possible attainment distributions with upper bound U by $\mathscr{X} := \cup_n \mathscr{X}_n$. The arithmetic mean function evaluated at any $\mathbf{x} \in \mathscr{X}$ is denoted by $\mu(\mathbf{x})$. Furthermore, let $\mathscr{X}_n^{\mu(\mathbf{x})}$ be the set of all attainment distributions of size n with upper bound U and *with the same mean as any* $\mathbf{x} \in \mathscr{X}_n$, and $\mathscr{X}^{\mu(\mathbf{x})}$ be the set of all possible attainment distributions with upper bound U and *with the same mean as any* $\mathbf{x} \in \mathscr{X}$.

Now, the following notation will be useful for studying minimum and maximum possible inequality within our framework. We denote the attainment distribution comprising n ones by $\mathbf{1}_n$, hence for any $\lambda \geq 0$, $\lambda \mathbf{1}_n$ is the constant or egalitarian distribution where all n elements are equal to λ. Next, for some $n \in \mathbb{N} \setminus \{1\}$ and for some $U \in \mathbb{R}_{++}$, let $\mathbb{G}_n = \{U/n, 2U/n, \ldots, (n-1)U/n\}$ denote a set of $n-1$ equally-spaced grid points between U/n and $(n-1)U/n$. A distribution $\mathbf{x} \in \mathscr{X}_n$ is *bipolar* whenever for some $n' \in \mathbb{N}$ such that $n' < n$, n' units in $\mathbf{x}$ attain the value of U and the remaining $n - n'$ units attain the value of 0. Clearly, since $n' \in \{1, ..., n-1\}$, for any bipolar distribution $\mathbf{x} \in \mathscr{X}_n$, $\mu(\mathbf{x}) \in \mathbb{G}_n$. Likewise, we refer to a distribution $\mathbf{x} \in \mathscr{X}_n$ as *almost-bipolar* whenever for some $n' \in \mathbb{N}$ such that $n' < n$, n' units in $\mathbf{x}$ attain the value of U, $n - n' - 1$ units in $\mathbf{x}$ attain the value of 0, and the leftover unit attains a value of $\varepsilon = [n\mu(\mathbf{x}) - n'U] \in (0,U)$. Accordingly, let $\mathscr{A} \subset \mathscr{X}$ be the set of all possible almost-bipolar distributions; $\mathscr{B} \subset \mathscr{X}$ be the set of all possible bipolar distributions; and let $\mathscr{M} = \mathscr{A} \cup \mathscr{B}$ be the set of all distributions that are either bipolar or almost bipolar. We assign subscript n to denote the subsets with population size n, i.e., $\mathscr{A}_n$, $\mathscr{B}_n$, and $\mathscr{M}_n$. Likewise, we use superscript $\mu(\mathbf{x})$ to denote the subsets with the same mean as $\mu(\mathbf{x})$, i.e., $\mathscr{A}^{\mu(\mathbf{x})}$, $\mathscr{B}^{\mu(\mathbf{x})}$, and $\mathscr{M}^{\mu(\mathbf{x})}$. Finally, we assign both n and $\mu(\mathbf{x})$ to denote the subsets with population size n and the same mean as $\mu(\mathbf{x})$, i.e., $\mathscr{A}_n^{\mu(\mathbf{x})}$, $\mathscr{B}_n^{\mu(\mathbf{x})}$, and $\mathscr{M}_n^{\mu(\mathbf{x})}$.

An *inequality index* $I : \mathscr{X} \to \mathbb{R}_+$ is a continuous real-valued function and is expected to satisfy certain basic properties (Chakravarty, 2009). The first basic property, *anonymity*, requires that an inequality index should not depend on a reordering of attainments across units. Formally, anonymity requires that $I(\mathbf{y}) = I(\mathbf{x})$ for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$ whenever $\mathbf{y} = \mathbf{xP}$, where $\mathbf{P}$ is a permutation matrix.[4] The second basic property, *transfer principle*, requires that a transfer from a richer to a poorer unit, without altering their relative positions, should decrease inequality (*progressive transfer*); whereas, alternatively, a transfer from a poorer to a richer unit should increase inequality (*regressive transfer*).[5] Formally, for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$, $I(\mathbf{y}) < I(\mathbf{x})$ whenever $\mathbf{y}$ is obtained from $\mathbf{x}$ by a progressive transfer and $I(\mathbf{y}) > I(\mathbf{x})$ whenever $\mathbf{y}$ is obtained from $\mathbf{x}$ by a regressive transfer.[6]

[4] A *permutation matrix* is a square matrix with exactly one element in each row and column equal to 1 and the rest of the elements are equal to zero.

[5] Technically, for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$, $\mathbf{y}$ is obtained from $\mathbf{x}$ by a *progressive transfer* whenever there are two units i, j and some $k > 0$ such that $y_i = x_i + k \leq x_j - k = y_j$ and $y_l = x_l$ for every $l \neq i, j$. Alternatively, $\mathbf{y}$ is obtained from $\mathbf{x}$ by a *regressive transfer* whenever there are two units i, j and some $k > 0$ such that $y_i + k = x_i \leq x_j = y_j - k$ and $y_l = x_l$ for every $l \neq i, j$.

[6] Some of the bounded indicators discussed in this paper are not literally transferable. For instance, we do not consider 'uneducating' highly educated individuals and transferring that education to less educated ones. Yet, one can compare two hypothetical scenarios, e.g. pre- and post-"progressive transfers", and still judge the latter exhibiting lower

Finally, we refer to a real valued function $f : \mathscr{X} \rightarrow \mathbb{R}_+$ as *symmetric* whenever $f(\mathbf{x}) = f(\mathbf{xP})$ for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$, where $\mathbf{P}$ is a permutation matrix. We refer to a real valued function $f : \mathscr{X} \rightarrow \mathbb{R}_+$ as *strictly S-convex* if, for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$, $f(\mathbf{y}) < f(\mathbf{x})$ whenever $\mathbf{y}$ is obtained from $\mathbf{x}$ by a progressive transfer and $f(\mathbf{y}) > f(\mathbf{x})$ whenever $\mathbf{y}$ is obtained from $\mathbf{x}$ by a regressive transfer (Marshall and Olkin, 1979, p. 53-54).

3 The consistency problem

The *consistency* problem has received significant attention in the literature on inequality measurement with bounded variables. Some variables such as mortality or literacy rates, access to basic facilities, etc., can be represented by either their distance from the lower bound (i.e. as attainments) or, alternatively, by their distance from the upper bound (i.e. as shortfalls). If $\mathbf{x} \in \mathscr{X}_n$ denotes the attainment distribution then we define the *shortfall distribution* associated with it as $\mathbf{x}^S = (x_1^S, \ldots, x_n^S) \in \mathscr{X}_n$ with $x_i^S = U - x_i$ representing i's shortfall from the upper bound U. Since there is no a priori reason to prefer one representation (attainment or shortfall) over the other, the literature has proposed different versions of consistency properties that an inequality ordering should satisfy.

The first consistency property, *perfect complementarity*, requires that the value of the inequality index remains unaltered when we switch between attainment and shortfall representations of the same distribution, i.e., $I(\mathbf{x}) = I(\mathbf{x}^S)$ for any $\mathbf{x} \in \mathscr{X}_n$ (Erreygers, 2009). The second one, *strong consistency*, requires that inequality measures should rank pairs of attainment distributions and their shortfall counterparts in a coherent manner. In other words, the inequality ranking should be robust to alternative representations of the variable., i.e., $I(\mathbf{x}) \leq I(\mathbf{y}) \Leftrightarrow I(\mathbf{x}^S) \leq I(\mathbf{y}^S)$ for any $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$ (Lambert and Zheng, 2011). Clearly, perfect complementarity implies strong consistency, but the reverse is not true. The third consistency property, *weak consistency* (Bosmans, 2016), is predicated on the realisation that it is possible to find pairs of different inequality indices that produce consistent comparisons as long as one index I^A is used for the attainment distribution and another index $I^S = \phi(I^A)$ is used for the shortfall counterpart, where $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a strictly increasing function. The pair (I^A, I^S) is *jointly* weakly consistent if and only if $I^A(\mathbf{x}) \leq I^A(\mathbf{y}) \Leftrightarrow I^S(\mathbf{x}^S) \leq I^S(\mathbf{y}^S)$ for any $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$. For example, if $I^A(\mathbf{x})$ is the Gini coefficient evaluated on the attainment distribution $\mathbf{x}$, then $I^S(\mathbf{x}^S) = \mu(\mathbf{x}^S) I^A(\mathbf{x}^S)/[U - \mu(\mathbf{x}^S)]$ provides a jointly weakly consistent inequality evaluation for the shortfall distribution.

Some advocates of perfect complementarity and strong consistency (Erreygers, 2009; Lambert and Zheng, 2011; Chakravarty et al., 2015; Seth and Alkire, 2017) suggest using absolute inequality indices (and related partial orderings).[7] However, Lasso de la Vega and Aristondo (2012) showed that strong consistency is satisfied by a wider class of indices derived from equally weighted generalised means of any inequality index evaluated at the attainment distribution and the same index evaluated at the corresponding shortfall distribution. Finally, Bosmans (2016) showed that weak consistency is satisfied by a broad class of *pairs of inequality indices*, including relative ones like the Gini coefficient and its respective weakly-consistent counterpart.[8]

inequality than the former.

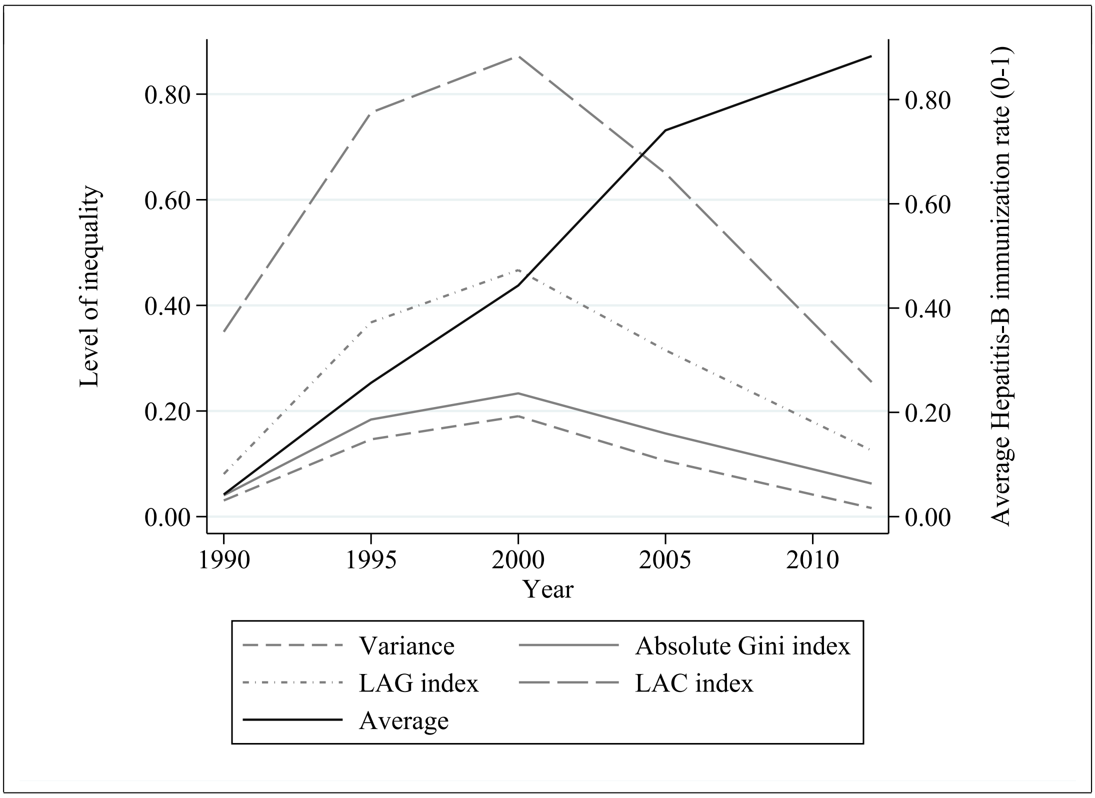
[7]The value of an *absolute* inequality index remains unchanged when all attainments are added the same amount, i.e., if it satisfies the *translation invariance* property, whereby $I(\mathbf{y}) = I(\mathbf{x})$ for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$ whenever $\mathbf{y} = \mathbf{x} + \lambda \mathbf{1}_n$ for some $\lambda \in \mathbb{R}$.

[8]The value of a *relative* inequality index remains unchanged when all attainments are altered in the same proportion, i.e., if it satisfies the *scale invariance* property, whereby $I(\mathbf{y}) = I(\mathbf{x})$ for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$ whenever $\mathbf{y} = \lambda \mathbf{x}$ for some $\lambda > 0$.

4 The boundary problem

Compared to the consistency problem, the *boundary problem* has received little attention. Generally, the boundary problem stems from the existence of a predictable functional relation between mean attainment and maximum feasible inequality (for a given inequality index), which hinders comparisons between distributions of bounded variables with different means. Indeed, when means differ it can be hard to disentangle the change in inequality due to socioeconomic phenomena (arguably the one we are interested in) from that due to the predicable relationship between mean attainment and maximum feasible inequality. For instance, in the case of many absolute inequality measures like the absolute Gini and the variance (which are consistent) when mean attainment converges toward any of the bounds, maximum feasible inequality falls prompting observed inequality to decrease in turn. In fact, with these absolute inequality measures the relationship between mean attainment and maximum feasible inequality usually follows an 'inverted U' shape with a global maximum at $\mu(\mathbf{x}) = 0.5U$ (if the lower bound is equal to 0).

Figure 1: Change in the mean and different absolute inequality measures for cross-country Hepatitis-B immunization rates



Source: Authors' own computations.

Notes: The formula for the variance is $V(\mathbf{x}) = (\sum_{i=1}^{n}[x_i - \mu(\mathbf{x})]^2)/n$. The formula for the absolute Gini index is $G_a(\mathbf{x}) = \mu(\mathbf{x})G_r(\mathbf{x})$ where G_r is the relative Gini index: $G_r(\mathbf{x}) = (\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|)/2\mu(\mathbf{x})n^2$. The LAG index is equal to $LAG(\mathbf{x}) = ([G_r(\mathbf{x})^{-1} + G_r(\mathbf{x}^S)^{-1}]/2)^{-1}$. The LAC index is equal to $LAC(\mathbf{x}) = ([CV(\mathbf{x})^{-1} + CV(\mathbf{x}^S)^{-1}]/2)^{-1}$, where $CV(\mathbf{x}) \equiv \sqrt{V(\mathbf{x})}/\mu(\mathbf{x})$ is the coefficient of variation.

In order to illustrate the empirical relevance of the boundary problem, figure 1 shows the change in mean Hepatitis-B immunization rates across countries as well as the changes in various measures of cross-country inequality between 1990 and 2012. Data for 150 countries are available at the UNdata website.[9] We use four different consistent inequality measures: two absolute inequality indices—the Variance V and the absolute Gini coefficient G_a (relative Gini coefficient *times* the mean); together with two inequality measures from the class of measures proposed by Lasso de la Vega and Aristondo (2012) (see note in figure 1 for details). As mean attainment increases from around to 4% in 1990 to around 88% in 2012, the values of all four inequality measures first increase and then decrease.

Thus, we observe a *Kuznets curve* relating mean attainment to each of the four consistent inequality measures over time. Is this type of relationship merely an empirical regularity or a mechanical artefact driven by the susceptibility of existing inequality indices (consistent or not) to the boundary problem?

In order to explore whether there is a predictable relationship between mean attainment and maximum possible inequality in the distribution of a bounded variable, first we identify whether such distributions reflecting maximum inequality for a given mean exist and, if so, what they look like. Then we need to check whether inequality indices, evaluated at such distributions, change in predictable ways when the mean is different.

Proposition 1 establishes the existence of a set of *maximum-inequality distributions* (MIDs) and shows that the set of MIDs associated to any distribution $\mathbf{x} \in \mathscr{X}_n$ is in fact equal to $\mathscr{M}_n^{\mu(\mathbf{x})} = \mathscr{X}_n^{\mu(\mathbf{x})} \cap \mathscr{M}$. Based exclusively on the transfer principle and the anonymity property, such MIDs are defined as the distributions that maximize inequality among all possible distributions with the *same population size n* and the *same mean* $\mu(\mathbf{x})$. Formally, first we need to define, for every $\mathbf{x} \in \mathscr{X}_n$, a partially ordered set $(\mathscr{X}_n^{\mu(\mathbf{x})}, \succeq)$ such that for any pair $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n^{\mu(\mathbf{x})}$, (1) $\mathbf{y} \succ \mathbf{x}$, which reads "**y** is more unequal than **x**", if **y** is obtained from **x** through a sequence of regressive transfers with or without additional permutations; and (2) $\mathbf{y} \sim \mathbf{x}$, which reads "**y** is as unequal as **x**" if **y** is obtained from **x** only through a sequence of permutations. Then we can state proposition 1:

Proposition 1 For any $n \in \mathbb{N} \setminus \{1\}$ and for any $\mathbf{x} \in \mathscr{X}_n$ such that $\mu(\mathbf{x}) \in (0, U)$, a set of maximum inequality distributions $\mathscr{M}_n^{\mu(\mathbf{x})} = \mathscr{X}_n^{\mu(\mathbf{x})} \cap \mathscr{M}$ constituting the maximal elements of the partially ordered set $(\mathscr{X}_n^{\mu(\mathbf{x})}, \succeq)$ exists and the elements of $\mathscr{M}_n^{\mu(\mathbf{x})}$ are *bipolar* when $\mu(\mathbf{x}) \in \mathbb{G}_n$ or *almost-bipolar* when $\mu(\mathbf{x}) \notin \mathbb{G}_n$.

Proof. See Appendix A1. ■

As it turns out, the MIDs are either bipolar or almost bipolar. Bipolar distributions consist of units with values at either the lower bound or upper bound exclusively, with at least one unit at each bound (as otherwise, should all units have the same values, the distribution would be egalitarian). Meanwhile, almost bipolar distributions consist of all units with either the lower or upper bound value, except for one unit with an interior value of $\varepsilon \in (0, U)$. The elements included in $\mathscr{M}_n^{\mu(\mathbf{x})}$ are unique up to permutations, that is: given any two elements $\mathbf{x}, \mathbf{y} \in \mathscr{M}_n^{\mu(\mathbf{x})}$, then $\mathbf{y} = \mathbf{xP}$ for some permutation matrix $\mathbf{P}$. Finally, also note that a set of MIDs defined in proposition 1 is unique for a given mean and for a given n. In other words, for any two distributions $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n^{\mu(\mathbf{x})}$ such that $\mathbf{x} \neq \mathbf{y}$, $\mathscr{M}_n^{\mu(\mathbf{x})} = \mathscr{M}_n^{\mu(\mathbf{y})}$.

Example 1: Assume $U = 1$, $n = 4$ and so $\mathbb{G}_4 = \{0.25, 0.5, 0.75\}$. Consider the distribution $\mathbf{x} = (0.1, 0.4, 0.7, 0.8)$ with $\mu(\mathbf{x}) = 0.5 \in \mathbb{G}_4$. In this case, the corresponding set of MIDs $\mathscr{M}_4^{\mu(\mathbf{x})}$ contains all possible permutations of the distribution $\hat{\mathbf{x}} = (0, 0, 1, 1)$, which is bipolar, and clearly $\mu(\hat{\mathbf{x}}) = \mu(\mathbf{x}) = 0.5$. Now consider a second distribution $\mathbf{y} = (0.2, 0.4, 0.7, 0.9)$ with $\mu(\mathbf{y}) = 0.55 \notin \mathbb{G}_4$. The corresponding MIDs $\mathscr{M}_4^{\mu(\mathbf{y})}$, in

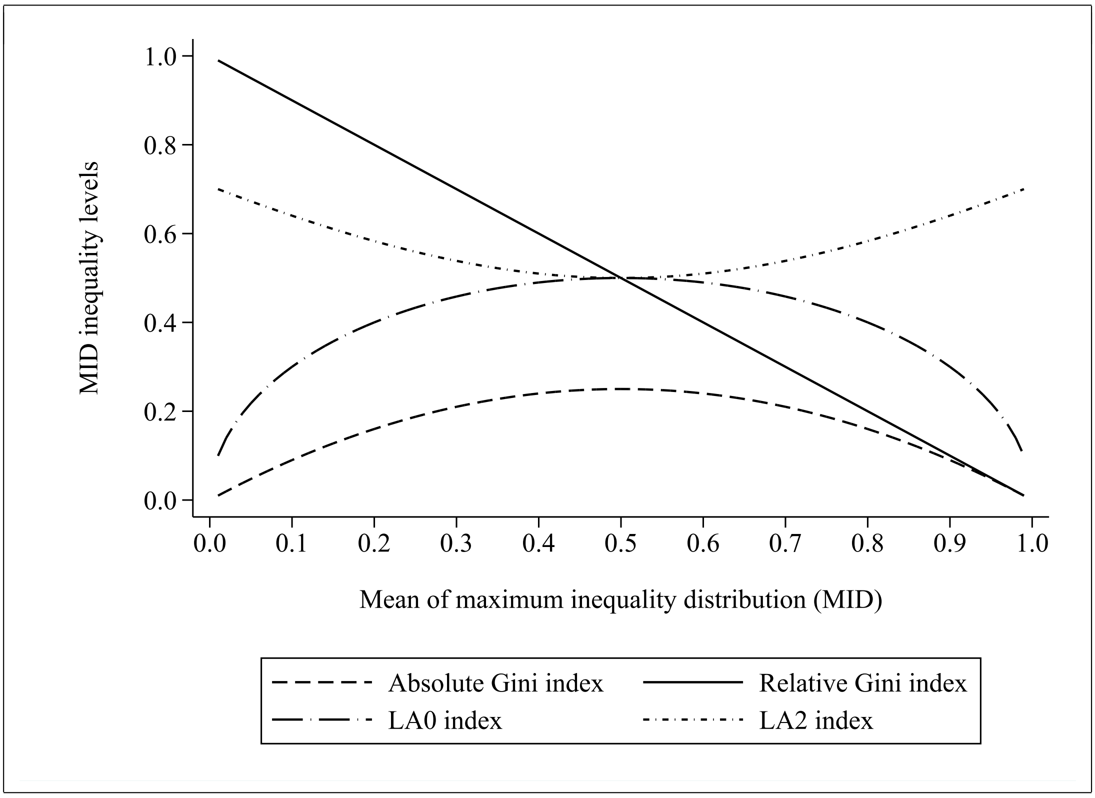
[9] https://data.un.org/Data.aspx?q=Hepatitis&d=WHO&f=MEASURE_CODE%3aWHS4_117.

this case, are all possible permutations of the distribution $\hat{\mathbf{y}} = (0, 0.2, 1, 1)$ with $\mu(\hat{\mathbf{y}}) = \mu(\mathbf{y}) = 0.55$ and $\varepsilon = 0.2 \in (0,1)$, but $\hat{\mathbf{y}}$ is almost-bipolar and no further regressive transfers are possible.

Even though MIDs are hypothetical distributions unlikely to be observed in practice, they do represent the benchmark case of maximum inequality against which we can compare distributions of bounded variables. The latter's inequality evaluations cannot be larger than their MID's as long as an inequality index I satisfies anonymity and the transfer principle.

Figure 2 shows the relationship between the mean (on the horizontal axis) and the level of inequality for the corresponding MIDs (vertical axis) using four different inequality measures that satisfy some form of consistency. For this exercise, we assume that $U = 1$ and $0 < \mu(\mathbf{x}) < 1$ (when $\mu(\mathbf{x}) \in \{0,1\}$, $\mathbf{x}$ is an egalitarian distribution). For simplicity, we assume that the MIDs used to calculate maximal inequality are bipolar (see further details in subsection 5.2). We choose the strongly-consistent absolute Gini index (Lambert and Zheng, 2011), the weakly-consistent relative Gini index (Bosmans, 2016), and two strongly-consistent indices from the class proposed by Lasso de la Vega and Aristondo (2012): the equally weighted geometric mean (LA0) and Euclidean mean (LA2), respectively, of the relative Gini indices evaluated at attainment and corresponding shortfall distributions.

Figure 2: The relationship between the mean and the maximum inequality of a bounded variable



Source: Authors' own computations.

Notes: The figure is based on $U = 1$ and bipolar distributions $\mathbf{x} \in \mathscr{B}_n^{\mu(\mathbf{x})}$. When evaluated at an MID $\mathbf{x} \in \mathscr{B}_n^{\mu(\mathbf{x})}$, the absolute Gini index equals $G_a(\mathbf{x}) = \mu(\mathbf{x})[1-\mu(\mathbf{x})]$ and the relative Gini index equals $G_r(\mathbf{x}) = 1-\mu(\mathbf{x})$. LA0 and LA2 are indices from the class proposed by Lasso de la Vega and Aristondo (2012). LA0 is the geometric mean of the (relative) Gini index evaluated at $\mathbf{x}$ and the same index evaluated at $\mathbf{x}^S$, whereas LA2 is the Euclidean mean of those two indices. Thus $LA0(\mathbf{x}) = \sqrt{\mu(\mathbf{x})[1-\mu(\mathbf{x})]}$ and $LA2(\mathbf{x}) = \sqrt{[\mu(\mathbf{x})^2 + (1-\mu(\mathbf{x}))^2]/2}$.

How do the selected inequality measures react as the mean increases along the horizontal axis in figure 2 and we move from one MID to another? Interestingly, the absolute Gini index and the LA0 index first increase, reach their maximum values at $\mu(\mathbf{x}) = 0.5$ and then decrease. By contrast, the LA2 index first decreases, reaches its minimum value around $\mu(\mathbf{x}) = 0.5$ and then increases. Finally, the relative Gini index decreases monotonically as the mean increases.

Evidently, inequality comparisons of distributions with *different means* using existing inequality indices suffer from the boundary problem, unless we control for changes in maximum possible inequality as a function of the mean. In order to control for the effect of differing means, we propose a simple approach: for a given distribution and any inequality index, we suggest comparing the observed inequality level against the corresponding MID's.

5 The class of normalized inequality indices

Our proposal to solve the boundary problem requires measuring observed inequality as a proportion of the maximum inequality attainable with a distribution having the same mean attainment, i.e. evaluated at the respective MID. The proposal restores comparability between distributions with different means consistently across alternative representations. In section 5.1, we characterize the class of normalized inequality indices assuming fixed population sizes. Then, section 5.2 presents the class of normalized inequality indices for varying population sizes. Finally, section 5.3 elucidates key differences between the two classes of normalized inequality indices, as well as vis-a-vis non-normalized counterparts, resorting to iso-inequality contours for Gini indices.

5.1 Normalized indices for fixed population

The crucial property imposed to counter the boundary problem is the *maximality principle*:

Maximality Principle For any $\mathbf{x} \in \mathscr{X}_n$, $I(\mathbf{x}) = 1$ whenever $\mathbf{x} \in \mathscr{M}_n^{\mu(\mathbf{x})}$.

This is a normalization property requiring that (i) inequality levels can never exceed the value of 1, and (ii) maximal inequality is attained only whenever $\mathbf{x}$ is an MID according to proposition 1 (wherein no further regressive transfers are feasible). As shown below, the maximality principle ensures that inequality is measured as a proportion of the maximum level reachable given a mean attainment so that any distribution different from the corresponding MID obtains an inequality value strictly below 1.

We also introduce the standard *equality principle*, which requires that $I(\mathbf{x}) = 0$ whenever $\mathbf{x} = \lambda \mathbf{1}_n$ for any $\mathbf{x} \in \mathscr{X}_n$ and $\lambda \geq 0$. This property ensures that inequality is minimal and equal to zero whenever all units feature exactly the same value for the indicator, i.e. $x_1 = \cdots = x_n$. From now onwards, we refer to the family of inequality indices satisfying the maximality principle and the equality principle in addition to other desirable properties as the class of *normalized inequality indices*, which is characterised in theorem 1:

Theorem 1 For any $n \in \mathbb{N} \setminus \{1\}$ and any $\mathbf{x} \in \mathscr{X}_n$, an inequality index I satisfies anonymity, the transfer principle, the equality principle, the maximality principle and strong consistency if and only if

$$I(\mathbf{x}) = \begin{cases} \dfrac{f(\mathbf{x}) - f(\bar{\mathbf{x}})}{f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}})} & \text{for } \mathbf{x} \neq \bar{\mathbf{x}} \\ 0 & \text{for } \mathbf{x} = \bar{\mathbf{x}} \end{cases}, \tag{1}$$

where $\bar{\mathbf{x}} = \mu(\mathbf{x})\mathbf{1}_n$, $\hat{\mathbf{x}} \in \mathscr{M}_n^{\mu(\mathbf{x})}$, $f : \mathscr{X}_n \to \mathbb{R}_{++}$ is a symmetric and strictly S-convex function, and $I(\mathbf{x}) = (g \circ I)(\mathbf{x}^S)$ where g is a strictly increasing function.

Proof. See Appendix A2. ■

According to theorem 1, a normalized inequality index $I(\mathbf{x})$ in our proposed class evaluated at distribution $\mathbf{x}$ is equal to any symmetric and S-convex function $f(\mathbf{x})$ evaluated at $\mathbf{x}$, subtracted by its corresponding minimum possible value $f(\bar{\mathbf{x}})$ evaluated at $\bar{\mathbf{x}}$, and then normalized by the difference between its corresponding maximum possible value $f(\hat{\mathbf{x}})$ evaluated at any of its uniquely associated MIDs, namely $\hat{\mathbf{x}} \in \mathscr{M}_n^{\mu(\mathbf{x})}$, and its corresponding minimum possible value $f(\bar{\mathbf{x}})$ evaluated at $\bar{\mathbf{x}}$.

All indices in our proposed class conveniently range between zero and one, where the former value is achieved in the absence of inequality and the latter corresponds to an MID. The value of a normalized index increases with a regressive transfer and decreases owing to a progressive transfer. A normalized index satisfies strong consistency as long as the index evaluated at an attainment distribution is a strictly increasing function of the *same* index evaluated at its shortfall counterpart, i.e., $I(\mathbf{x}) = (g \circ I)(\mathbf{x}^S)$.

Crucially, numerous functional forms of f are admissible, including (but certainly not limited to) different classes of relative, absolute, intermediate, super-relative or super-absolute inequality measures.[10] For instance, all absolute inequality indices including the variance, the absolute Gini or the members of the class proposed by Chakravarty et al. (2015) are admissible forms of f. Likewise, some relative indices like the (relative) Gini index and the coefficient of variation are admissible because they can be decomposed into the product of an absolute inequality index times a function of the mean.[11]

Examples

We present a few examples of normalized inequality indices derived from popular inequality measures that will be applied in the empirical section 6. For convenience of presentation in this section, we refer to the normalised inequality index corresponding to the admissible form f as f^* (instead of I), in order to clarify that f^* is derived from an admissible f. That is, f^* is the normalized version of f.

When f is the absolute or the relative Gini index (i.e., $f(\mathbf{x}) = G_a(\mathbf{x})$ or $f(\mathbf{x}) = G_r(\mathbf{x})$) then it is easy to check (see Appendix A5) that

$$G_a^*(\mathbf{x}) = G_r^*(\mathbf{x}) = \begin{cases} \dfrac{G_a(\mathbf{x})U}{\mu(\mathbf{x})(U-\mu(\mathbf{x}))} & \text{if } \mathbf{x} \neq \bar{\mathbf{x}} \text{ and } \mu(\mathbf{x}) \in \mathbb{G}_n \\ \dfrac{G_a(\mathbf{x})n^2}{(n-n'-1)\varepsilon + n'(U-\varepsilon) + n'(n-n'-1)U} & \text{if } \mathbf{x} \neq \bar{\mathbf{x}} \text{ and } \mu(\mathbf{x}) \notin \mathbb{G}_n \ .^{12} \\ 0 & \text{if } \mathbf{x} = \bar{\mathbf{x}} \end{cases} \quad (2)$$

Thus, $G_a^*(\mathbf{x})$ compares $G_a(\mathbf{x})$ against the maximal inequality value that such index could possibly take for any distribution with mean equal to $\mu(\mathbf{x})$ (which equals $\mu(\mathbf{x})(U-\mu(\mathbf{x}))/U$ when the MID is bipolar, i.e.

[10] See Bosmans (2016) for a concise and comprehensive typology.

[11] The admissible forms of f also include the class of indices proposed by Lasso de la Vega and Aristondo (2012) as it is straightforward to verify that for all indices in their class $f(\mathbf{x}) = f(\mathbf{x}^S)$ for all $\mathbf{x}$ and thus $I(\mathbf{x}) = I(\mathbf{x}^S)$ for all $\mathbf{x}$.

[12] The reader is reminded that $n' < n$ is the number of units in $\mathbf{x}$ attaining U and $\varepsilon = [n\mu(\mathbf{x}) - n'U]$ (section 2).

when $\mu(\mathbf{x}) \in \mathbb{G}_n$). Remarkably, the normalized inequality indices derived from the absolute and the relative Gini indices coincide. To simplify notation, such normalized Gini index will be referred to as $G^*(\mathbf{x})$.

We can also derive the normalized versions of the standard deviation ($f(\mathbf{x}) = \sigma(\mathbf{x}) = \sqrt{V(\mathbf{x})}$) and the coefficient of variation ($f(\mathbf{x}) = CV(\mathbf{x}) = \sigma(\mathbf{x})/\mu(\mathbf{x})$). It is easy to check (see Appendix A5) that

$$\sigma^*(\mathbf{x}) = CV^*(\mathbf{x}) = \begin{cases} \dfrac{\sigma(\mathbf{x})}{\sqrt{\mu(\mathbf{x})(U-\mu(\mathbf{x}))}} & \text{if } \mathbf{x} \neq \bar{\mathbf{x}} \text{ and } \mu(\mathbf{x}) \in \mathbb{G}_n \\ \dfrac{\sigma(\mathbf{x})\sqrt{n}}{\sqrt{(n-n'-1)\mu(\mathbf{x})^2 + n'(U-\mu(\mathbf{x}))^2 + (\varepsilon-\mu(\mathbf{x}))^2}} & \text{if } \mathbf{x} \neq \bar{\mathbf{x}} \text{ and } \mu(\mathbf{x}) \notin \mathbb{G}_n \\ 0 & \text{if } \mathbf{x} = \bar{\mathbf{x}} \end{cases}. \quad (3)$$

Once again, the normalised version of an absolute inequality index and its relative counterpart coincide. More generally, whenever $a(\mathbf{x})$ is an absolute inequality index and $r(\mathbf{x}) = a(\mathbf{x})/\mu(\mathbf{x})$ is its relative counterpart, one can easily check that $a^*(\mathbf{x}) = r^*(\mathbf{x})$ for all $\mathbf{x} \in \mathscr{X}_n$.

While all normalized inequality measures are defined in the same way (i.e., as a fraction of the maximal inequality level that can be reached given a mean attainment and a fixed population size; see Theorem 1), their explicit mathematical representation differs slightly depending on whether $\mu(\mathbf{x}) \in \mathbb{G}_n$ or $\mu(\mathbf{x}) \notin \mathbb{G}_n$. It is easy to check that, when the decimal precision one is working with is fixed (i.e., numbers are represented with a precision of $k \in \mathbb{N}$ decimals, so all numbers with a higher number of decimals are rounded), then for sufficiently large values of n (more specifically, when $n \geq 10^k/U$) one always has $\mu(\mathbf{x}) \in \mathbb{G}_n$. When this happens, the corresponding MID is bipolar, thus leading to more compact formulations for the corresponding normalized inequality measures and easier calculations (see equations 2 and 3). For instance, in the fairly common case where $k = 2$ and $U = 1$, when $n \geq 100$ then all MIDs are bipolar (as in our empirical illustration in section 6).

5.2 Comparing populations with different sizes

At least since Dalton (1920), the most popular answer to the challenge of comparing inequality across distributions with different population sizes is the *population principle*, which requires that identical cloning of all units should leave inequality unaltered (thereby rendering populations with different sizes comparable).[13] Formally, the population principle requires that whenever $\mathbf{y}$ is obtained from $\mathbf{x}$ by a *replication* for some $\mathbf{x}, \mathbf{y} \in \mathscr{X}$, then $I(\mathbf{y}) = I(\mathbf{x})$; where $\mathbf{y} \in \mathscr{X}_{n'}$ for some $n' = \gamma n$ and $\gamma \in \mathbb{N} \setminus \{1\}$ is said to be obtained from $\mathbf{x} \in \mathscr{X}_n$ by a *replication*, whenever $\mathbf{y} = (\mathbf{x}, \ldots, \mathbf{x})$, i.e. γ copies of $\mathbf{x}$ are repeated one after the other in $\mathbf{y}$. A normalized inequality measure from the class defined in theorem 1 does not comply with the population principle even when an admissible functional form of f does, because even though the replication of a bipolar MID is itself an MID, the replication of an almost bipolar MID is *not* an MID, based on how proposition 1 defines an MID.[14]

Therefore if we want our normalized inequality measures to fulfil the population principle we must adopt

[13] For a more general proposal, see Aboudi et al. (2010).

[14] For instance, when $n = 2$ and $U = 1$, an MID associated to a distribution with mean equal to 0.25 is $(0, 0.5)$. However, the replication $(0, 0.5, 0, 0.5)$ of that MID is not an MID itself. The corresponding MID for a distribution with $n = 4$ and with mean equal to 0.25 is in fact $(0, 0, 0, 1)$.

a different definition of the set of MIDs, one compliant with the population principle. In this section, we assume that the attainment values take rational number (i.e., $x_i \in [0,U] \cap \mathbb{Q}$).[15] We denote the set of all attainment distributions, where attainments take rational numbers, by $\mathscr{Q}$ and that with the same mean as $\mathbf{x} \in \mathscr{Q}$ by $\mathscr{Q}^{\mu(\mathbf{x})}$. Proposition 2 establishes the existence of a set of MIDs and shows that the set of MIDs, associated with all distributions *sharing the same mean* across *all population sizes* is, in this case, equal to $\mathscr{B}^{\mu(\mathbf{x})} = \mathscr{Q}^{\mu(\mathbf{x})} \cap \mathscr{B}$. Based on the transfer and population principles combined with anonymity, these MIDs are defined as the distributions that maximize inequality among all possible distributions with the *same mean* but *varying population sizes*. Formally, we now need to define a different partially ordered set $(\mathscr{Q}^{\mu(\mathbf{x})}, \succeq)$ such that for any pair $\mathbf{x}, \mathbf{y}$: (1) $\mathbf{y} \succ \mathbf{x}$, which reads "$\mathbf{y}$ is more unequal than $\mathbf{x}$", if $\mathbf{y}$ is obtained from $\mathbf{x}$ through a sequence of regressive transfers with or without additional permutations and/or replications; and (2) $\mathbf{y} \sim \mathbf{x}$, which reads "$\mathbf{y}$ is as unequal as $\mathbf{x}$" if $\mathbf{y}$ is obtained from $\mathbf{x}$ only through a sequence of permutations and/or replications. Then, we can state proposition 2:

Proposition 2 For any $\mathbf{x} \in \mathscr{Q}$ such that $\mu(\mathbf{x}) \in (0,U)$, a set of maximum inequality distributions $\mathscr{B}^{\mu(\mathbf{x})} = \mathscr{Q}^{\mu(\mathbf{x})} \cap \mathscr{B}$ constituting the maximal elements of the partially ordered set $(\mathscr{Q}^{\mu(\mathbf{x})}, \succ)$ exists and all elements of $\mathscr{B}^{\mu(\mathbf{x})}$ are *bipolar*.

Proof. See Appendix A3. ■

According to proposition 2, in a setting compliant with the population principle, only bipolar distributions maximize inequality. Thus, the Maximality Principle introduced in section 5 must be adapted and rewritten as follows:

Restricted Maximality Principle For any $\mathbf{x} \in \mathscr{Q}$, $I(\mathbf{x}) = 1$ whenever $\mathbf{x} \in \mathscr{B}^{\mu(\mathbf{x})}$.

With this reformulated version of the Maximality Principle, we can now axiomatically characterise the class of normalized inequality indices compliant with the population principle:

Theorem 2 For any $n \in \mathbb{N} \setminus \{1\}$ and for any $\mathbf{x} \in \mathscr{Q}_n$, an inequality index I satisfies anonymity, the transfer principle, the equality principle, the restricted maximality principle, strong consistency and *the population principle* if and only if

$$I(\mathbf{x}) = \begin{cases} \dfrac{f(\mathbf{x}) - f(\bar{\mathbf{x}})}{f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}})} & \text{for } \mathbf{x} \neq \bar{\mathbf{x}} \\ 0 & \text{for } \mathbf{x} = \bar{\mathbf{x}} \end{cases},$$

where $\bar{\mathbf{x}} = \mu(\mathbf{x})\mathbf{1}_n$, $\hat{\mathbf{x}} \in \mathscr{B}^{\mu(\mathbf{x})}$, $f : \mathscr{Q} \to \mathbb{R}_{++}$ is a symmetric and strictly S-convex function satisfying the *population principle*, and $I(\mathbf{x}) = (g \circ I)(\mathbf{x}^S)$ where g is a strictly increasing function.

Proof. See Appendix A4. ■

Theorem 2 implies that we can construct normalized inequality indices that abide by the population principle as long as the inequality index for both numerator and denominator (i.e., f) satisfies the population principle and *we evaluate the chosen inequality index at any bipolar distribution with mean equal to $\mu(\mathbf{x})$ in the denominator of $I(\mathbf{x})$*. The ensuing normalized inequality indices satisfy the population principle alongside all the properties listed in theorem 1.

[15]The rationality requirement is apparent in the proof of Proposition 2, see Appendix A3. For practical applications this restriction is inconsequential as it is difficult to conceive reasonable situations where attainment levels should be irrational numbers.

5.3 Comparing different approaches

We have characterized two types of normalized inequality indices: those applicable to fixed population sizes (section 5.1) and those suitable for variable population sizes satisfying the population principle (section 5.2). Drawing on a simple example, this sub-section explains how these indices relate to each other, as well as to standard inequality measures proposed in the literature.

The two approaches to measuring normalised inequality (corresponding to the two definitions of MIDs and their respective classes of indices) bear a large degree of overlap. In fact, the formulae for normalized inequality indices compliant with the population principle (section 5.2) is identical to the corresponding formulae for indices suitable for fixed population sizes (section 5.1) whenever $\mu(\mathbf{x}) \in \mathbb{G}_n$.[16] As argued before, when the population size n is sufficiently large and the decimal precision is kept fixed (as is the case in our empirical illustration as well as in several applications in practice), the condition $\mu(\mathbf{x}) \in \mathbb{G}_n$ is always satisfied.

At the other extreme, as n gets smaller, it becomes increasingly unlikely that $\mu(\mathbf{x}) \in \mathbb{G}_n$, so the occurrence of almost bipolar MIDs represents the rule rather than the exception. Indeed, in the extreme case when $n = 2, \mathbb{G}_2$ only has one element $(U/2)$ and the MIDs for all distributions are almost bipolar, except for those with mean equal to $U/2$. In this context, it is illustrative to show how the different normalized inequality measures behave and compare *vis-à-vis* each other, and with respect to standard inequality measures for the simplest non-trivial case (i.e., for $n = 2$). For that purpose, Figure 3 presents the iso-inequality contours of the absolute Gini index ($G_a(x_1,x_2)$, panel A), the relative Gini index ($G_r(x_1,x_2)$, panel B), the normalized Gini index based on Theorem 1 (i.e., for fixed population; $G^*(x_1,x_2)$, panel C), and the normalized Gini index complying with the Population Principle ($G^*_P(x_1,x_2)$, panel D), in the case where $U = 1$ (Appendix A6 shows how we arrive at these iso-inequality contours).[17]

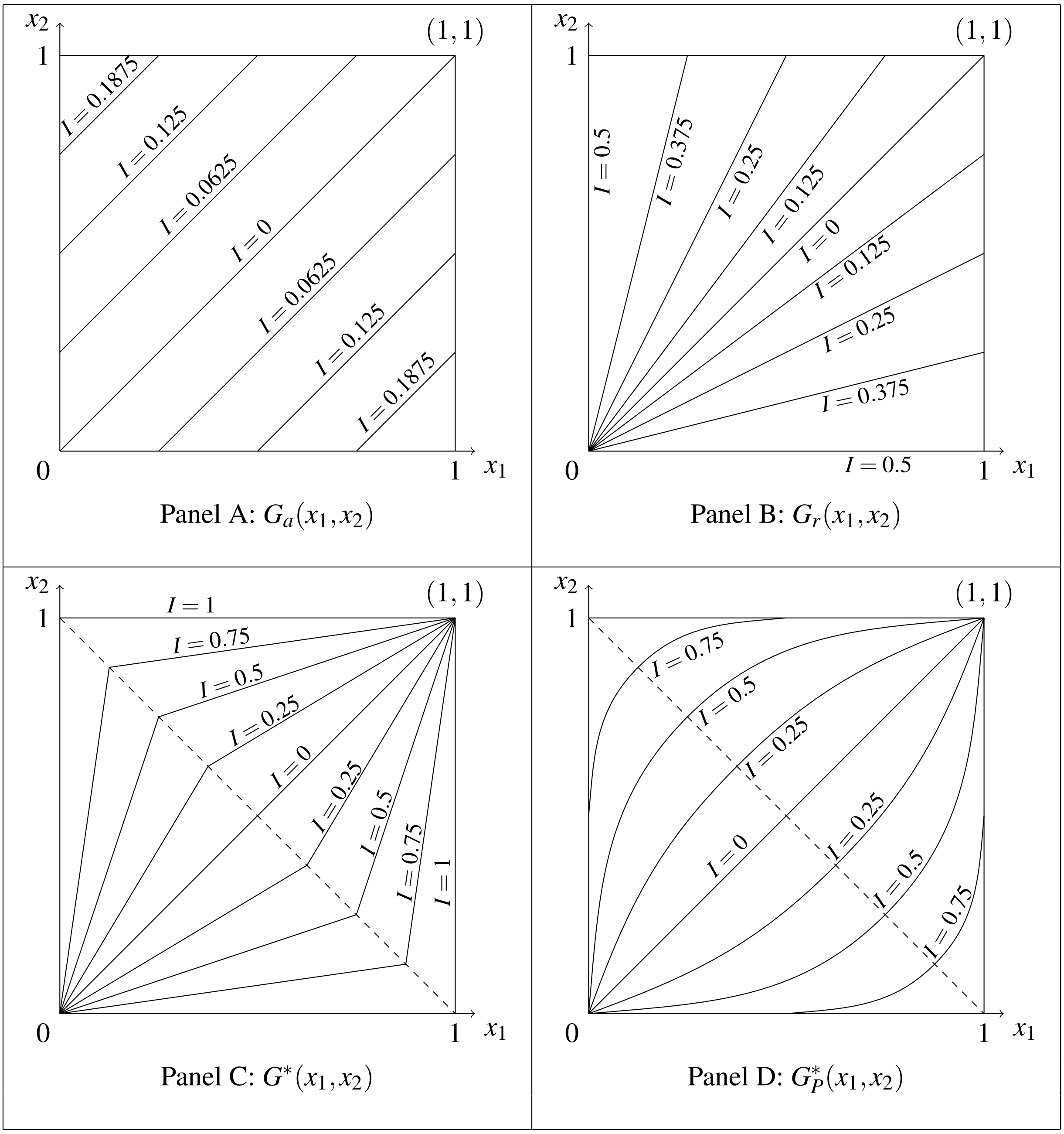
As is well-known, $G_a(x_1,x_2) \in [0,0.25]$ and the iso-inequality contours for $G_a(x_1,x_2)$ are parallel to the 45° line, while $G_r(x_1,x_2) \in [0,0.5]$ and the iso-inequality contours for $G_r(x_1,x_2)$ are straight lines 'emanating from' (or 'converging to') the origin $(0,0)$. In contrast, the iso-inequality contours for the two normalised Gini indices, $G^*(x_1,x_2) \in [0,1]$ and $G^*_P(x_1,x_2) \in [0,1]$, exhibit completely different shapes. In the case of $G^*(x_1,x_2)$, all level contours are made of two line segments meeting in the diagonal $\{(x_1,x_2) \in [0,1]^2 \mid x_1 + x_2 = 1\}$, which, together, connect the points $(0,0)$ and $(1,1)$. Their shapes (though not their corresponding inequality levels) coincide with the level contours of $G_r(x_1,x_2)$ when $\mu(x_1,x_2) \leq 1/2$ and with those of $G_r(x_1^S,x_2^S)$ when $\mu(x_1,x_2) \geq 1/2$ (where $x_1^S = 1-x_1$ and $x_2^S = 1-x_2$, see Appendix A6). In addition, one has that $G^*(x_1,x_2) = G^*(x_1^S,x_2^S)$. Lastly, the level contours $G^*_P(x_1,x_2) = c$ (where $c \in [0,1]$) are curves that (i) are symmetrical with respect to the $x_2 = 1-x_1$ axis for all $c \in [0,1]$ (i.e., $G^*_P(x_1,x_2) = G^*_P(x_1^S,x_2^S)$), and (ii) they connect the points $(0,0)$ and $(1,1)$ when $c \leq 1/2$.

As can be inferred from panel C, all the distributions (x_1,x_2) lying at the border of the unit square maximize inequality (i.e., they are MIDs) when the latter is measured with $G^*(x_1,x_2)$. By contrast, panel D shows that, when the population principle is imposed, only the bipolar distributions, namely $(0,1)$ and $(1,0)$, maximize inequality. The variegated shapes of the iso-inequality contours when moving from one inequality measure to another explains the discrepancies that might exist among them – a phenomenon that is clearly illustrated in the empirical section that follows.

[16]Readers are reminded of our examples of normalized inequality indices in section 5.1, whose formulae vary depending on whether $\mu(\mathbf{x}) \notin \mathbb{G}_n$ (i.e., almost bipolar MIDs) or $\mu(\mathbf{x}) \in \mathbb{G}_n$ (i.e., bipolar MIDs).

[17]Results remain essentially unaltered when the absolute and relative Gini indices are substituted by the standard deviation and the coefficient of variation, respectively.

Figure 3: Iso-inequality contours for the different Gini coefficient ($n = 2$ and $U = 1$)

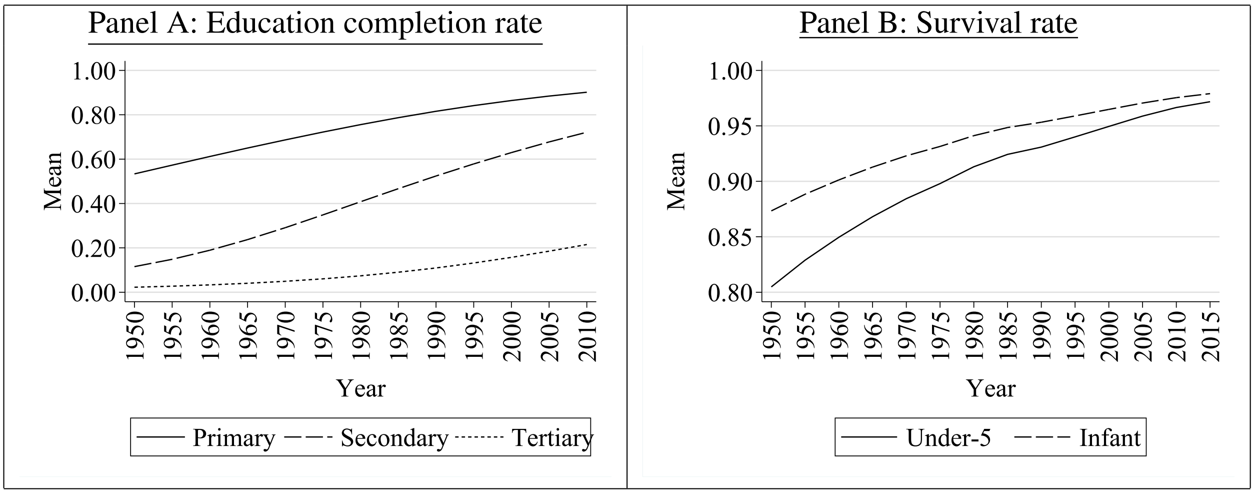


Notes: The figure is based on $n = 2$ and $U = 1$. $G_a(x_1, x_2)$ in panel A is the absolute Gini index applied to an attainment distribution. $G_r(x_1, x_2)$ in panel B is the relative Gini index applied to the same attainment distribution. $G^*(x_1, x_2)$ in panel C is the normalised Gini index applied to the attainment distribution for fixed population. $G^*_P(x_1, x_2)$ in panel in panel D is the normalised Gini index applied to the attainment distribution for variable population.

6 Empirical illustration: Evolution of cross-country inequality in education and health

In order to illustrate the empirical relevance of our proposal, we study the evolution of cross-country inequality in three education indicators and two health indicators since 1950. To examine inequality trends in education, we look at the share of total adult population with at least some primary education, the share of total adult population with at least some secondary education and the share of total adult population with at least some tertiary education. For health inequality we select the under-five survival rate and the infant survival rate, namely the attainment complements of the respective mortality rates. The education data come from the Barro-Lee dataset; whereas, the data on mortality rates were obtained from the United Nations' Department of Economic and Social Welfare website.[18] The three education indicators are rates, thus bounded between zero and one. For the health indicators, we obtain the survival rates by subtracting the mortality rates from 1,000 and then normalising the differences by 1,000. Therefore, the survival rates also lie between zero and one. Education and health indicators' data are available for 133 and 201 countries, respectively.[19]

Figure 4: Changes in the means of selected education indicators and health indicators since 1950



Source: Authors' own computations.

Figure 4 shows the change in mean attainments for the education indicators between 1950 and 2010 and for the health indicators between 1950 and 2015, for every five-year period.[20] All global averages display steady improvements since 1950. Between 1950 and 2010, the mean for the primary completion rate increases from 0.53 to 0.92, the mean for the secondary completion rate rises from 0.12 to 0.76, and the tertiary completion rate increases from 0.02 to 0.25. Meanwhile, the mean under-5 survival rate increases from 0.80 in 1950 to 0.97 in 2015 and the mean infant survival rate increases from 0.87 in 1950 to 0.98 in 2015.

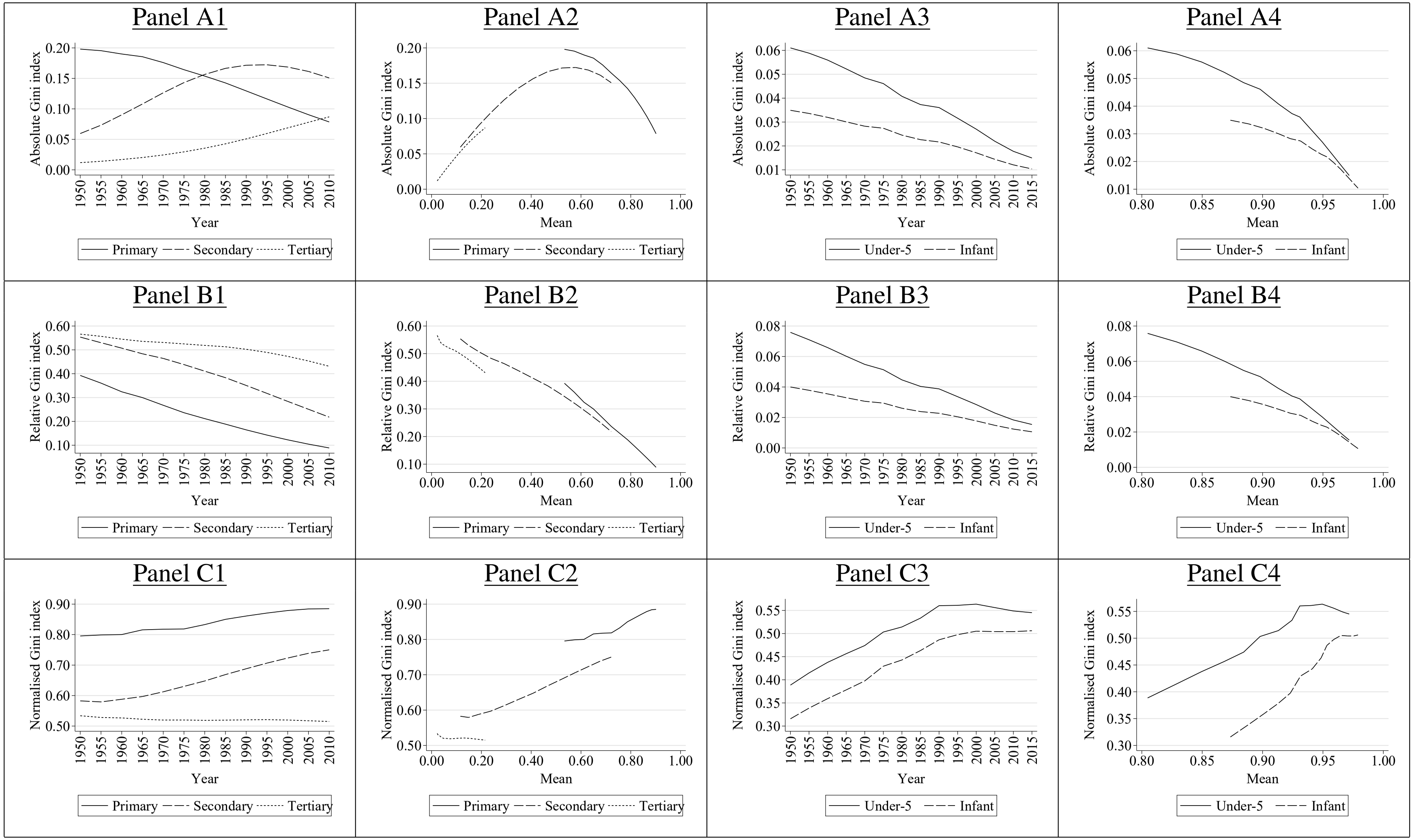
Panels A1 and A3 in Figure 5 present the trends in absolute Gini since 1950 for all five indicators. Interestingly, absolute cross-country inequality in primary completion rate falls throughout, which is compatible with the mean attainment in the indicator steadily increasing from above 0.5 since 1950. Similarly, absolute inequality in the tertiary completion rate rises throughout, which is compatible with the mean attainment in the indicator increasing from very low levels since 1950 and still remaining below 0.5. By contrast, inequality in the secondary completion rates rises between 1950 and around 1990, then falling thereafter, which is

[18] Source of the education data: http://www.barrolee.com/. Source of health data: https://population.un.org/wpp/Download/Standard/Mortality/.

[19] Since we report results with a precision of $k = 2$ decimals, and the variables we are dealing with have an upper bound of $U = 1$, the number of observations (n) is large enough to ensure that $\mathbf{x} \in \mathbb{G}_n$ for any $\mathbf{x} \in \mathscr{X}_n$, so we can use the bipolar MIDs when calculating normalized inequality levels.

[20] The mean attainment does not include population weights. Each country, irrespective of its size, is considered as a unit with equal importance.

Figure 5: Change in cross-country inequality according to absolute, relative and normalised Gini indices for education and health indicators since 1950



Source: Authors’ own computations.

also compatible with the mean of the indicator surpassing 0.50 around that year. The trends of absolute Gini indices for the two health indicators are the same as those of inequality in the primary completion rate as the means lie above 0.50 for the entire period. The relations between mean attainments and absolute Gini indices are presented in Panels A2 and A4, where each horizontal axis represents mean attainment and the vertical axis displays the absolute Gini index. The seeming Kuznets-curve relationship between the mean and absolute Gini index appears to be quite predictable and 'mechanical' in our illustration.

Panels B1-B4 in Figure 5 show the evolution of relative inequality measured by the relative Gini index. In contrast to the different absolute Gini patterns across indicators, relative Gini indices fall throughout as mean attainments increase steadily for all five indicators (Panels B1 and B3). The relationships between the mean attainments and relative Gini indices are depicted in Panels B2 and B4. Both absolute and relative Gini indices follow the pattern predicted for their maximum values in Section 4.

Finally, panels C1–C4 present the trends in normalized Gini indices, G^*.[21] Panel C1 shows that, unlike both absolute and relative Gini indices, the normalized Gini indices register an upward trend for primary completion rates. Therefore, after we control for the boundary problem, cross-country inequality in this education indicator does not appear to have fallen, contrary to the trends generated by the absolute and relative Gini indices. The normalized Gini indices for the secondary completion rate follow the same upward trend between 1950 and 2010, but the normalized Gini indices for the tertiary completion rate linger between 0.51 and 0.54 for the entire period. Meanwhile, panel C3 shows that the normalized Gini indices for the under-5 survival rate increase until 2000 and then fall. Similarly, the normalized Gini indices for the infant survival rate increase until 2000, but then stabilise afterwards. Therefore, our empirical illustration clearly documents that normalized Gini indices may produce very different inequality trends from their traditional absolute and relative counterparts.

7 Concluding remarks

The use of standard income inequality measures to study the dispersion of bounded variables poses several problems. On the one hand, the bounded domain of the variables generate a predictable functional relation between the mean of a distribution and its maximum inequality levels. For instance, in the case of many indices, chiefly absolute ones, when the mean approaches either the upper or the lower bound, inequality mechanically goes to zero. On the other hand, several measures of inequality, e.g. all the relative ones, fail to rank distributions consistently between the alternatives of attainments and shortfall representations. We proposed a new approach to inequality measurement aimed at solving both problems, consistency and boundary, *simultaneously*. Basically, the proposed normalized inequality measures compare observed inequality levels against the maximum inequality level achievable with the same measure across all hypothetical distributions having the same mean.[22] The normalized inequality indices are strongly consistent (Bosmans, 2016) and eliminate any mechanical relationship between mean attainment and maximum inequality brought about by the boundary problem.

As mentioned in the introduction, several solutions have been proposed for the consistency problem, including reliance on absolute Lorenz curves and indices (Lambert and Zheng, 2011), generalised means of indices evaluated at both attainment and shortfall distribution (Lasso de la Vega and Aristondo, 2012) and a relaxation of the (strong) consistency requirement, partially (Bosmans, 2016) or completely (Kenny, 2004). Remarkably, our solution represents an alternative to all the aforementioned. By contrast, to the best of our

[21]Recall that G^* can be obtained by normalizing either the absolute or the relative Gini index.
[22]Or same mean and population size when the population principle is not upheld.

knowledge, many fewer alternative proposals to ours exist to handle the boundary effect. One alternative to the class of normalized inequality indices may suggest taking the natural logarithm of the bounded variable, thereby eliminating the lower bound when its value is zero. However, this approach is plagued with serious problems. To begin with, it would not work when an untransformed value is equal to zero. Meanwhile, if both bounds were positive, taking logarithms for all values would be feasible, but would not solve the boundary effect (because the two bounds would just be replaced by new bounds). Moreover, if the upper bound were higher than 1 then logarithmic transformations would compress the dispersion of values between 1 and the upper bound while expanding it for values between 0 and 1, with the concomitant ambiguous effect on inequality rankings. Worse still, some inequality measures like the variance applied to logarithms of a variable are known to violate the popular transfer principle (Foster and Ok, 1999). Finally, inequality comparisons based on logarithmic transformations of bounded variables would be generally inconsistent. In a nutshell, the costs and inconveniences of the logarithmic transformation render it an unappealing alternative.

To illustrate the empirical relevance of our methodological proposal, we investigated the extent of cross-country inequality in selected health and education indicators since 1950, comparing the absolute and relative Gini indices against their unique normalized counterpart. The normalized Gini index portrays a dramatically different picture of global dissemination in human development achievements. The absolute and relative Gini indices point to cross-country inequality reductions in several health and education indicators, but these trends look suspicious in the sense that they are highly correlated with mean attainment values. More specifically, both absolute and relative Gini-measured inequality decrease as mean attainment approaches its upper bound. This concern is confirmed by the normalized Gini index, whose trends suggest that global progress in the studied health and education indicators did not follow more conceivably egalitarian paths.

While our proposed inequality measures successfully address the consistency and boundary problems, there remain other measurement challenges in the context of bounded variables. For example, Lasso de la Vega and Aristondo (2012) provide conditions whose fulfillment guarantees robustness of inequality comparisons to changes in the upper bound. Though admittedly this problem is not that serious when bounds are neither arbitrary nor expected to change across time and space (e.g. in the case of indicators expressed as percentage ratios), it is nonetheless worth exploring how our proposed measurement framework could accommodate this potential concern. This challenge is left for future research. Likewise, future research could explore partial orderings respecting the consistency and normalization properties that were combined to generate the normalized inequality indices.

References

Aboudi, R., D. Thon, and S. Wallace (2010). Inequality comparisons when the populations differ in size. *Journal of Economic Inequality 8*, 47–70.

Atkinson, A. (1970). On the measurement of inequality. *Journal of Economic Theory 2*, 244–63.

Atkinson, A. (2018). *Inequality. What can be done?* Cambridge, Massachusetts: Harvard University Press.

Bosmans, K. (2016). Consistent comparisons of attainment and shortfall inequality: A critical examination. *Health Economics 25*(11), 1425–32.

Bourguignon, F. (2017). *The globalization of inequality*. New Jersey: Princeton University Press.

Chakravarty, S. (2009). *Inequality, polarization and poverty. Advances in distributional analysis*. Springer.

Chakravarty, S., N. Chattopadhyay, and C. D'Ambrosio (2015). On a family of achievement and shortfall inequality indices. *Health Economics 25*(12), 1503–1513.

Clarke, P., U. Gerdtham, M. Johannesson, K. Bingefors, and L. Smith (2002). On the measurement of relative and absolute income-related health inequality. *Social Science and Medicine 55*, 1923–8.

Dalton, H. (1920). The measurement of the inequality of incomes. *The Economic Journal 30*(119), 348–361.

Erreygers, G. (2009). Can a single indicator measure both attainment and shortfall inequality? *Journal of Health Economics 28*, 885–93.

Erreygers, G. and T. Van Ourti (2011). Measuring socioeconomic inequality in health, health case and health financing by means of rank-dependent indices: A recipe for good practice. *Journal of Health Economics 30*, 685–94.

Foster, J. and E. Ok (1999). Lorenz dominance and the variance of logarithms. *Econometrica 67*(4), 901–907.

Kenny, C. (2004). Why are we worried about income? Nearly everything that matters is converging. *World Development 33*(1), 1–19.

Lambert, P. and B. Zheng (2011). On the consistent measurement of attainment and shortfall inequality. *Journal of Health Economics 30*, 214–9.

Lasso de la Vega, C. and O. Aristondo (2012). Proposing indicators to measure achievement and shortfall inequality consistently. *Journal of Health Economics 31*, 578–583.

Marshall, A. W. and I. Olkin (1979). *Inequalities: Theory of majorization and its applications*. New York: Academic Press.

Micklewright, J. and K. Stewart (1999). Is the well-being of children converging in the European Union? *Economic Journal 109*(459), 692–714.

Milanovic, B. (2018). *Global inequality. A new approach for the age of globalization*. Cambridge, Massachusetts: Harvard University Press.

Piketty, T. (2015). *The economics of inequality*. Cambridge, Massachusetts: Harvard University Press.

Prados de la Escosura, L. (2021). Augmented human development in the age of globalization. *The Economic History Review 74*(4), 946–975.

Seth, S. and S. Alkire (2017). Did poverty reduction reach the poorest of the poor? Complementary measures of poverty and inequality in the counting approach. In S. Bandyopadhyay (Ed.), *Research on Economic Inequality*, Volume 25, pp. 63–102. Emerald Publishing Limited.

Seth, S. and M. McGillivray (2018). Composite indices, alternative weights, and comparison robustness. *Social Choice and Welfare 51*, 657–679.

Wagstaff, A. (2005). The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Economics 14*, 429–32.

Wagstaff, A. (2009). Correcting the concentration index: A comment. *Journal of Health Economics 28*, 516–20.

Wagstaff, A. (2011). The concentration of a binary outcome revisited. *Health Economics 20*(10), 1155–60.

Appendices

Appendix A1 Proof of Proposition 1

Let us start with an $\mathbf{x} \in \mathscr{X}_n \setminus \mathscr{M}$ for some $n \in \mathbb{N} \setminus \{1\}$ such that $\mu(\mathbf{x}) \in (0, U)$ (i.e., $\mathbf{x}$ is neither bipolar nor almost bipolar). Given that in the proposition's partial order $(\mathscr{X}_n^{\mu(\mathbf{x})}, \succeq)$, a regressive transfer increases inequality while a permutation keeps it unaltered and both keep the mean unaltered, we may always perform a sequence of regressive transfers (with or without additional permutations) until exhaustion to obtain any element of $\mathscr{M}$ that belongs in the set of distributions with the same population size and the same mean, namely $\mathscr{X}_n^{\mu(\mathbf{x})}$.[23]

Now, there can be two types of cases: (i) $\mu(\mathbf{x}) \in \mathbb{G}_n$ and (ii) $\mu(\mathbf{x}) \notin \mathbb{G}_n$, where $\mathbb{G}_n = \{U/n, \ldots, (n-1)U/n\}$ is the set of $n-1$ equally-spaced grid points between 0 and U.

Case (i): Whenever $\mu(\mathbf{x}) \in \mathbb{G}_n$, then there exists a natural number $n' \leq n$ such that $\mu(\mathbf{x}) = n'U/n$. Clearly, $\mu(\mathbf{x}) = n' \times U/n + (n-n') \times 0/n$. Starting with $\mathbf{x}$, it is possible to have a series of regressive transfers until a distribution with n' elements equalling U and $n-n'$ elements equalling zero is reached. In this case, the set of MIDs is $\mathscr{X}_n^{\mu(\mathbf{x})} \cap \mathscr{B}$.

Case (ii): Whenever $\mu(\mathbf{x}) \notin \mathbb{G}_n$, then there exist a natural number $n' \leq n$ such that $n'U/n < \mu(\mathbf{x}) < (n'+1)U/n$. In this case, a series of regressive transfers are possible until n' elements are equal to U and $n-n'-1$ elements are equal to zero. Note that it is not possible for $n'+1$ elements to be equal to U because $\mu(\mathbf{x}) < (n'+1)U/n$. However, when n' elements are equal to U, then $\mu(\mathbf{x}) > n'U/n$ and therefore the remaining element will have a value of $\varepsilon = n\mu(\mathbf{x}) - n'U$ so that: $\mu(\mathbf{x}) = n' \times U/n + (n-n'-1) \times 0/n + \varepsilon/n$. It is straightforward to verify that $\varepsilon \in (0, U)$. In this case, the set of MIDs is $\mathscr{X}_n^{\mu(\mathbf{x})} \cap \mathscr{A}$.

Thus, the maximum inequality distribution (MID) for $\mathbf{x}$ is an element in the set $\mathscr{X}_n^{\mu(\mathbf{x})} \cap (\mathscr{A} \cup \mathscr{B}) = \mathscr{X}_n^{\mu(\mathbf{x})} \cap \mathscr{M}$, which by our definition is equal to $\mathscr{M}_n^{\mu(\mathbf{x})}$. Now, whenever $\mathbf{x} \in \mathscr{X}_n \cap \mathscr{M}$ for some $n \in \mathbb{N} \setminus \{1\}$ (i.e., $\mathbf{x}$ is either bipolar or almost bipolar), it can be trivially checked that $\mathbf{x} \in \mathscr{M}_n^{\mu(\mathbf{x})}$. Hence, a set of MIDs for any $\mathbf{x} \in \mathscr{X}_n$ such that $\mu(\mathbf{x}) \in (0, U)$ always exists and constitutes the set of maximal elements $\mathscr{M}_n^{\mu(\mathbf{x})}$ of the partially ordered set $(\mathscr{X}_n^{\mu(\mathbf{x})}, \succeq)$. ■

Appendix A2 Proof of theorem 1

We first prove the *sufficiency* part. Consider some $\mathbf{x} \in \mathscr{X}_n$ for some $n \in \mathbb{N} \setminus \{1\}$ and so the set of corresponding MIDs is $\mathscr{M}_n^{\mu(\mathbf{x})}$ by proposition 1. We then already know that

$$I(\mathbf{x}) = \begin{cases} \dfrac{f(\mathbf{x}) - f(\bar{\mathbf{x}})}{f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}})} & \text{for } \mathbf{x} \neq \bar{\mathbf{x}} \\ 0 & \text{for } \mathbf{x} = \bar{\mathbf{x}} \end{cases}, \tag{A1}$$

[23] Note that each element within $\mathbf{x}$ is bounded between 0 and U by definition and so it is not possible to perform further regressive transfers once the bounds are reached. The proof proceeds in similar line of argument as the proof of Theorem 1 in Seth and McGillivray (2018).

where $\bar{\mathbf{x}} = \mu(\mathbf{x})\mathbf{1}_n$, $\hat{\mathbf{x}} \in \mathscr{M}_n^{\mu(\mathbf{x})}$, $f : \mathscr{X}_n \to \mathbb{R}_{++}$ is a symmetric and strictly S-convex function, and $I(\mathbf{x}) = (g \circ I)(\mathbf{x}^S)$ where g is a strictly increasing function.

We now show that I satisfies the required properties. (*i*) Since $f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}}) > 0$ because any $\hat{\mathbf{x}} \in \mathscr{M}_n^{\mu(\mathbf{x})}$ can be obtained from $\bar{\mathbf{x}}$ by a series of regressive transfers and f is strictly S-convex, it follows directly from the formulation in Equation A1 that I satisfies the *equality principle* as $I(\bar{\mathbf{x}}) = 0$, and the *maximality principle* as $I(\hat{\mathbf{x}}) = 1$ for any $\hat{\mathbf{x}} \in \mathscr{M}_n^{\mu(\mathbf{x})}$. (*ii*) Suppose that $\mathbf{y} \in \mathscr{X}_n$ is obtained from $\mathbf{x}$ such that $\mathbf{y} = \mathbf{xP}$, where $\mathbf{P}$ is a permutation matrix. By definition, $\mu(\mathbf{x}) = \mu(\mathbf{y})$, $\mathscr{M}_n^{\mu(\mathbf{y})} = \mathscr{M}_n^{\mu(\mathbf{x})}$ and $\bar{\mathbf{y}} = \bar{\mathbf{x}}$. Provided f is symmetric, $f(\mathbf{y}) = f(\mathbf{x})$ and also $f(\hat{\mathbf{y}}) = f(\hat{\mathbf{x}})$ for any $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathscr{M}_n^{\mu(\mathbf{x})}$. So, $I(\mathbf{y}) = I(\mathbf{x})$. Thus I satisfies *anonymity*. (*iii*) Suppose $\mathbf{y}' \in \mathscr{X}_n$ is obtained from $\mathbf{x}$ by a regressive transfer. Again, by definition, $\mu(\mathbf{x}) = \mu(\mathbf{y}')$, $\mathscr{M}_n^{\mu(\mathbf{y}')} = \mathscr{M}_n^{\mu(\mathbf{x})}$ and $\bar{\mathbf{y}}' = \bar{\mathbf{x}}$. Provided f is strictly S-convex, $f(\mathbf{y}') > f(\mathbf{x})$ and so $I(\mathbf{y}') > I(\mathbf{x})$. Therefore, I satisfies the *transfer principle*. (*iv*) We show that I satisfies *strong consistency* by noting that $(g \circ I)(\mathbf{x}^S) \leq (g \circ I)(\mathbf{y}^S) \Leftrightarrow I(\mathbf{x}^S) \leq I(\mathbf{y}^S)$ for any $\mathbf{x}, \mathbf{y} \in \mathscr{X}_n$ since g is a strictly increasing function. Moreover, $(g \circ I)(\mathbf{x}^S) = I(\mathbf{x})$. Therefore $I(\mathbf{x}) \leq I(\mathbf{y}) \Leftrightarrow I(\mathbf{x}^S) \leq I(\mathbf{y}^S)$.

Let us now prove the *Necessity* part. Suppose I satisfies anonymity and the transfer principle. Then, I is symmetric and S-convex. Given that a monotonically increasing transformation of an S-convex function is also S-convex, we may write (without loss of generality) $I(\mathbf{x}) = af(\mathbf{x}) + b$ for some $\mathbf{x} \in \mathscr{X}_n$, where $a \in \mathbb{R}_{++}$, $b \in \mathbb{R}$ and f is some S-convex function. The equality principle requires that $I(\bar{\mathbf{x}}) = 0$, therefore $af(\bar{\mathbf{x}}) + b = 0$ or $b = -af(\bar{\mathbf{x}})$. On the other hand, the maximality principle requires that $I(\hat{\mathbf{x}}) = 1$ for any $\hat{\mathbf{x}} \in \mathscr{M}_n^{\mu(\mathbf{x})}$, thus $af(\hat{\mathbf{x}}) - af(\bar{\mathbf{x}}) = 1$ or $a = 1/[f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}})]$. Hence, inserting the values of a and b, we obtain:

$$I(\mathbf{x}) = \frac{f(\mathbf{x}) - f(\bar{\mathbf{x}})}{f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}})}. \tag{A2}$$

Finally, drawing on Bosmans (2016, proposition 4) we deduce that if I is strongly consistent then it must be the case that $I(\mathbf{x}) = (g \circ I)(\mathbf{x}^S)$ where g is a strictly increasing function. ■

Appendix A3 Proof of proposition 2

Proposition 1 holds for $\mathscr{X}$, hence it also holds for $\mathscr{Q} \subset \mathscr{X}$. Consider an $\mathbf{x} \in \mathscr{Q}_n \setminus \mathscr{M}$ for some $n \in \mathbb{N} \setminus \{1\}$ such that $\mu(\mathbf{x}) \in (0, U)$.[24] Then, with proposition 1 we can prove that the MIDs for $\mathbf{x}$ are the elements of $\mathscr{Q}_n^{\mu(\mathbf{x})} \cap \mathscr{M}$. Now there are two cases, where we will allow n to vary:

Case (i): First, suppose $\mu(\mathbf{x}) \in \mathbb{G}_n$ and $\mu(\mathbf{x}) = (i/n)U$ for some $i \in \mathbb{N}$ and $1 \leq i \leq n-1$. Then, according to proposition 1, the set of MIDs for $\mathbf{x}$ is $\mathscr{Q}_n^{\mu(\mathbf{x})} \cap \mathscr{B}$. Now, suppose for some $n' = \gamma n$ where $\gamma \in \mathbb{N} \setminus \{1\}$, $\mathbf{y} \in \mathscr{Q}_{n'}$ is obtained from $\mathbf{x} \in \mathscr{Q}_n$ by *replication*. Since $\mathbf{x}, \mathbf{y} \in \mathscr{Q}$, clearly $\mu(\mathbf{y}) = (i\gamma/n')U = (i/n)U = \mu(\mathbf{x})$. Again, by proposition 1, the set of MIDs for $\mathbf{y}$ is $\mathscr{Q}_{n'}^{\mu(\mathbf{x})} \cap \mathscr{B}$. Given that the results hold for any n and any γ, the set of MIDs for $\mathbf{x}$ is $\mathscr{Q}^{\mu(\mathbf{x})} \cap \mathscr{B}$.

Case (ii): Next, suppose $\mu(\mathbf{x}) \notin \mathbb{G}_n$. Then, according to proposition 1, the set of MIDs for $\mathbf{x}$ is $\mathscr{Q}_n^{\mu(\mathbf{x})} \cap \mathscr{A}$. Consider some $\hat{\mathbf{x}} \in \mathscr{Q}_n^{\mu(\mathbf{x})} \cap \mathscr{A}$, where $\hat{\mathbf{x}}$ has n_1 elements equal to zero, $n - n_1 - 1$ elements equal to U and one element equal to $\varepsilon \in (0, U)$. Further suppose for some $n' = \delta n$, where $\delta \in \mathbb{N} \setminus \{1\}$ and $\delta = \lambda U/\varepsilon$ for some $\lambda \in \mathbb{N}$, that $\mathbf{y} \in \mathscr{Q}_{n'}$ is obtained from $\mathbf{x} \in \mathscr{Q}_n$ by *replication* such that $\mu(\mathbf{y}) \in \mathbb{G}_{n'}$. Such replication

[24]Note that the cases $\mu(\mathbf{x}) = \{0, U\}$ are purposefully dismissed from proposition 2.

is always feasible since $\varepsilon \in \mathbb{Q}$. Clearly, $\mu(\mathbf{x}) = \mu(\mathbf{y})$ using the same line of argument in the preceding paragraph. Thus, the set of MIDs for $\mathbf{y}$ by proposition 1 is $\mathscr{Q}_{n'}^{\mu(\mathbf{x})} \cap \mathscr{B}$ since $\mu(\mathbf{y}) \in \mathbb{G}_{n'}$. We now need to show that an element in $\mathscr{Q}_{n'}^{\mu(\mathbf{x})} \cap \mathscr{B}$ is more unequal than an element in $\mathscr{Q}_{n}^{\mu(\mathbf{x})} \cap \mathscr{A}$ in the partially ordered set $(\mathscr{Q}^{\mu(\mathbf{x})}, \succeq)$. Let $\hat{\mathbf{y}}$ be obtained from $\hat{\mathbf{x}}$ by *replication* by the same δ times so that $\hat{\mathbf{y}}$ has δn_1 elements equal to zero, $\delta(n - n_1 - 1)$ elements equal to U and δ elements equal to ε. Based on the partially ordered set $(\mathscr{Q}^{\mu(\mathbf{x})}, \succeq)$, by the population principle, $\hat{\mathbf{y}} \sim \hat{\mathbf{x}}$. Clearly, further regressive transfers are possible in $\hat{\mathbf{y}}$ and a series of regressive transfers until exhaustion leads to some $\hat{\mathbf{z}} \in \mathscr{Q}_{n'}^{\mu(\mathbf{x})} \cap \mathscr{B}$ since, by definition, $\delta\varepsilon = \lambda U$. It follows that, in $(\mathscr{Q}^{\mu(\mathbf{x})}, \succeq)$, $\hat{\mathbf{z}} \succ \hat{\mathbf{y}}$ and therefore $\hat{\mathbf{z}} \succ \hat{\mathbf{x}}$. Given that our result holds for an arbitrary n and an arbitrary δ, the MIDs for $\mathbf{x}$, in this case, are also $\mathscr{Q}^{\mu(\mathbf{x})} \cap \mathscr{B}$.

Finally, suppose $\mathbf{x} \in \mathscr{Q}_n \cap \mathscr{M}$, which means that either $\mathbf{x} \in \mathscr{Q}_n \cap \mathscr{A}$ (i.e., almost bipolar) or $\mathbf{x} \in \mathscr{Q}_n \cap \mathscr{B}$ (i.e., bipolar). Following the same line of argument presented in the two previous paragraphs, it can be shown that the MIDs for $\mathbf{x}$ are again $\mathscr{Q}^{\mu(\mathbf{x})} \cap \mathscr{B}$.

Hence, the set of MIDs for any $\mathbf{x} \in \mathscr{Q}$ such that $\mu(\mathbf{x}) \in (0, U)$ exists and is equal to $\mathscr{Q}^{\mu(\mathbf{x})} \cap \mathscr{B}$, which by our definition is $\mathscr{B}^{\mu(\mathbf{x})}$ and constitute the maximal elements of $(\mathscr{Q}^{\mu(\mathbf{x})}, \succeq)$. ■

Appendix A4 Proof of theorem 2

We first prove the sufficiency part. Applying theorem 1, which holds for $\mathscr{Q}_n \subset \mathscr{X}_n$, since it holds for $\mathscr{X}_n$, we can show that I satisfies *anonymity*, the *transfer principle*, the *equality principle*, and *strong consistency*. Now, since $f(\hat{\mathbf{x}}) > f(\bar{\mathbf{x}})$ for any $\hat{\mathbf{x}} \in \mathscr{B}^{\mu(\mathbf{x})}$ because any $\hat{\mathbf{x}} \in \mathscr{B}^{\mu(\mathbf{x})}$ can be obtained from $\bar{\mathbf{x}}$ by a series of regressive transfers with or without combinations of replications and permutations, and f is strictly S-convex, we have $I(\hat{\mathbf{x}}) = 1$; that is, I satisfies the *restricted maximality principle*. Finally, we prove that I satisfies the population principle. Let $\mathbf{y}$ be obtained from $\mathbf{x} \in \mathscr{Q}_n$ through a replication. Then, by definition, $\mu(\mathbf{x}) = \mu(\mathbf{y})$ and so by proposition 2, $\mathscr{B}^{\mu(\mathbf{x})} = \mathscr{B}^{\mu(\mathbf{y})}$. It is also straightforward to verify that $\bar{\mathbf{y}}$ is a replication of $\bar{\mathbf{x}}$. Therefore, based on $(\mathscr{Q}^{\mu(\mathbf{x})}, \succeq)$, $\mathbf{y} \sim \mathbf{x}$ and $\bar{\mathbf{y}} \sim \bar{\mathbf{x}}$, and hence $f(\bar{\mathbf{y}}) = f(\bar{\mathbf{x}})$ and $f(\mathbf{y}) = f(\mathbf{x})$ because f satisfies the population principle. Coupled with $f(\hat{\mathbf{y}}) = f(\hat{\mathbf{x}})$ for any $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathscr{B}^{\mu(\mathbf{x})}$ and $f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}}) > 0$, clearly $I(\mathbf{y}) = I(\mathbf{x})$. Hence, I satisfies the population principle.

The proof of the necessity part is similar to the proof of theorem 1's necessity part, with some additional modifications. If I satisfies anonymity and the transfer principle, then, without loss of generality, $I(\mathbf{x}) = af(\mathbf{x}) + b$ for some $\mathbf{x} \in \mathscr{Q}_n$, where $a \in \mathbb{R}_{++}$, $b \in \mathbb{R}$ and f is some S-convex function. Recalling proposition 2, suppose for some $n' = \delta n$, where $\delta \in \mathbb{N} \setminus \{1\}$ and $\delta = \lambda U / \varepsilon$ for some $\lambda \in \mathbb{N}$, that $\mathbf{y} \in \mathscr{Q}_{n'}$ is obtained from $\mathbf{x}$ by *replication* such that $\mu(\mathbf{y}) \in \mathbb{G}_{n'}$. Given that I satisfies the population principle, then $I(\mathbf{y}) = I(\mathbf{x})$. It follows that $f(\mathbf{y}) = f(\mathbf{x})$ since $a > 0$ and so f also satisfies the population principle. By the equality principle, which requires that $I(\bar{\mathbf{x}}) = 0$ where $\bar{\mathbf{x}} = \mu(\mathbf{x})\mathbf{1}_n$, we obtain $af(\bar{\mathbf{x}}) + b = 0$ or $b = -af(\bar{\mathbf{x}})$. Now, suppose $\hat{\mathbf{x}}$ is obtained from $\mathbf{y}$ by a series of regressive transfers until exhaustion. Then, certainly $\hat{\mathbf{x}} \in \mathscr{B}^{\mu(\mathbf{y})}$ by proposition 2, since $\delta\varepsilon = \lambda U$ and $\mu(\mathbf{x}) = \mu(\mathbf{y})$ by construction. By the restricted maximality principle, we obtain $I(\hat{\mathbf{x}}) = 1$ and so $af(\hat{\mathbf{x}}) - af(\bar{\mathbf{x}}) = 1$ or $a = 1/[f(\hat{\mathbf{x}}) - f(\bar{\mathbf{x}})]$, which leads us to the normalized formulation. Finally, following the proof of theorem 1, it can be easily verified that if I satisfies strong consistency then $I(\mathbf{x}) = (g \circ I)(\mathbf{x}^S)$ where g is a strictly increasing function. ■

Appendix A5 Derivation of normalized inequality indices

Here we show the derivation of the inequality measures presented as examples in section 5.1. We start with the formulas relying on bipolar MIDs followed by formulas based on almost bipolar MIDs. The former are simpler than the latter.

In a bipolar MID $\hat{\mathbf{x}}$, assume that a share s of the population attains the value of 0 and the rest $(1-s)$ the value of U. Given that $\mu(\hat{\mathbf{x}}) = \mu(\mathbf{x})$, by definition, the following restriction must hold:

$$\mu(\hat{\mathbf{x}}) = s \times 0 + (1-s) \times U \Rightarrow (1-s)U = \mu(\mathbf{x}) \Rightarrow s = 1 - \frac{\mu(\mathbf{x})}{U}. \tag{A3}$$

The absolute Gini index:

Computing the absolute Gini index for $\hat{\mathbf{x}}$ yields:

$$G_a(\hat{\mathbf{x}}) = s(1-s)U. \tag{A4}$$

Plugging equation A3 into equation A4 and manipulating algebraically yields:

$$G_a(\hat{\mathbf{x}}) = \frac{\mu(\mathbf{x})}{U}\left(1 - \frac{\mu(\mathbf{x})}{U}\right)U = \frac{\mu(\mathbf{x})[U - \mu(\mathbf{x})]}{U}.$$

The relative Gini index:

Computing the relative Gini index for $\hat{\mathbf{x}}$ yields:

$$G_r(\hat{\mathbf{x}}) = \frac{s(1-s)U}{\mu(\mathbf{x})}. \tag{A5}$$

Plugging equation A3 into equation A5 and manipulating algebraically yields:

$$G_r(\hat{\mathbf{x}}) = \frac{U - \mu(\mathbf{x})}{U}.$$

The standard deviation:

Computing the standard deviation for $\hat{\mathbf{x}}$ yields:

$$\sigma(\hat{\mathbf{x}}) = \sqrt{s[\mu(\mathbf{x})]^2 + (1-s)[U - \mu(\mathbf{x})]^2}. \tag{A6}$$

Plugging equation A3 into equation A6 and manipulating algebraically, we obtain:

$$\sigma(\hat{\mathbf{x}}) = \sqrt{\left[1 - \frac{\mu(\mathbf{x})}{U}\right][\mu(\mathbf{x})]^2 + \frac{\mu(\mathbf{x})}{U}[U - \mu(\mathbf{x})]^2} = \sqrt{\mu(\mathbf{x})\,[U - \mu(\mathbf{x})]}.$$

The coefficient of variation:

Computing the coefficient of variation for $\hat{\mathbf{x}}$ yields:

$$CV(\hat{\mathbf{x}}) = \frac{\sigma(\hat{\mathbf{x}})}{\mu(\hat{\mathbf{x}})} = \frac{\sqrt{\mu(\mathbf{x})\,[U - \mu(\mathbf{x})]}}{\mu(\mathbf{x})} = \sqrt{\frac{U - \mu(\mathbf{x})}{\mu(\mathbf{x})}}. \tag{A7}$$

Derivation of normalised inequality indices for almost bipolar MIDs

In an almost bipolar MID $\hat{\mathbf{x}}$ we have n' units in the population with value U, one unit with value $0 < \varepsilon < U$ and the rest, $n - n' - 1$ with value 0. Moreover, $\varepsilon = n\mu(\mathbf{x}) - n'U$. For each of the denominators of the indices mentioned in section 5.1 we get:

The absolute Gini index:

$$G_a(\hat{\mathbf{x}}) = \frac{1}{2n^2}[(n-n'-1) \times 1 \times |0-\varepsilon| + (n') \times 1 \times |U-\varepsilon| + n'(n-n'-1) \times 1 \times |0-U|] \tag{A8}$$

Simplifying equation A8 we get the denominator of 2 for the almost bipolar case (noting later that the 2 in the fraction gets cancelled out as it also appears in the numerator's formula):

$$G_a(\hat{\mathbf{x}}) = \frac{1}{2n^2}[(n-n'-1)\varepsilon + (n')(U-\varepsilon) + n'(n-n'-1)U]. \tag{A9}$$

The relative Gini index:

Essentially we get the same formula for the denominator of 2 as in A8 but divided by $\mu(\mathbf{x})$ (again, the 2 in the fraction gets cancelled out as it also appears in the numerator's formula):

$$G_r(\hat{\mathbf{x}}) = \frac{1}{2n^2\mu(\mathbf{x})}[(n-n'-1)\varepsilon + (n')(U-\varepsilon) + n'(n-n'-1)U]. \tag{A10}$$

The standard deviation:

$$\sigma(\hat{\mathbf{x}}) = \sqrt{\frac{1}{n}[(n-n'-1)(0-\mu(\mathbf{x}))^2 + n'(U-\mu(\mathbf{x}))^2 + (\varepsilon-\mu(\mathbf{x}))^2]}. \tag{A11}$$

Simplifying equation A11 we get the denominator of 3 for the almost bipolar case.

The coefficient of variation:

We get the same formula as in A11 but divided by $\mu(\mathbf{x})$:

$$CV(\hat{\mathbf{x}}) = \frac{1}{\mu(\mathbf{x})}\sqrt{\frac{1}{n}[(n-n'-1)\mu(\mathbf{x})^2 + n'(U-\mu(\mathbf{x}))^2 + (\varepsilon-\mu(\mathbf{x}))^2]}. \tag{A12}$$

Finally, for each of the aforementioned indices (for bipolar and almost bipolar MIDs), we compute $f(\mathbf{x})/f(\hat{\mathbf{x}})$.

Appendix A6 Derivation of iso-inequality contours

When $n = 2$ and $U = 1$, the absolute Gini index can be written as $G_a(x_1,x_2) = |x_1 - x_2|/4$, where $(x_1,x_2) \in [0,1]^2$. In that case, the iso-inequality contours are straight lines parallel to the 45° line. In this setting, the relative Gini index can be written as $G_r(x_1,x_2) = |x_1 - x_2|/4\mu = |x_1 - x_2|/2(x_1 + x_2)$. Since this function is homogeneous of degree 0 (i.e., $G_r(\lambda x_1, \lambda x_2) = G_r(x_1,x_2)$ for all $\lambda > 0$), the iso-inequality contours are straight lines emanating from (or converging to) the origin $(0,0)$.

What about the iso-inequality contours for the normalized Gini index that does not comply with the population principle ($G^*(x_1,x_2)$)? Here, $\mathbb{G}_2 = \{1/2\}$, so there are basically two cases: either $\mu(x_1,x_2) \leq 1/2$

or $\mu(x_1,x_2) \geq 1/2$. Case (i): $\mu(x_1,x_2) = (x_1+x_2)/2 \leq 1/2$. Here, the MIDs associated with (x_1,x_2) are $\{(0,x_1+x_2),(x_1+x_2,0)\}$. When the absolute Gini index is applied to any of those distributions, one obtains $G_a(0,x_1+x_2) = G_a(x_1+x_2,0) = (x_1+x_2)/4$. Hence, $G^*(x_1,x_2) = |x_1-x_2|/(x_1+x_2)$. These are straight lines emanating from (or converging to) the origin $(0,0)$. Case (ii): $\mu(x_1,x_2) = (x_1+x_2)/2 \geq 1/2$. Now, the MIDs associated with (x_1,x_2) are $\{(x_1+x_2-1,1),(1,x_1+x_2-1)\}$. Calculating the absolute Gini index of any of those distributions yields $G_a(x_1+x_2-1,1) = G_a(1,x_1+x_2-1) = (2-(x_1+x_2))/4$. Hence $G^*(x_1,x_2) = |x_1-x_2|/(2-(x_1+x_2)) = |x_1^S - x_2^S|/(x_1^S+x_2^S)$ (where $x_1^S = 1-x_1$ and $x_2^S = 1-x_2$). These are straight lines emanating from (or converging to) the point $(1,1)$. Finally, it is easy to prove that the two sets of iso-inequality contours match at the intersection (i.e., for the set of points $\{(x_1,x_2) \in [0,1]^2 | x_1+x_2 = 1\}$, the values of the iso-inequality contours examined in cases (i) and (ii) coincide).

According to theorem 2 and equation 2, the normalized Gini index complying with the population principle is simply defined as $G_P^*(x_1,x_2) = G_a(x_1,x_2)/(\mu(x_1,x_2)(1-\mu(x_1,x_2)))$. Manipulating algebraically, one obtains that $G_P^*(x_1,x_2) = |x_1-x_2|/((x_1+x_2)(2-(x_1+x_2)))$. This is the function from which the iso-inequality contours shown in Figure 3 panel D are calculated.