Simple Distribution Sensitive Welfare Measures

Aart Kraay*[1] Christoph Lakner[1] Berk Özler[2] Benoit Decerf[1] Dean Jolliffe[1] Olivier Sterck[3] Nishant Yonzan[1]

November 4, 2025

Simple welfare measures such as mean income are ubiquitous but not distribution sensitive. In contrast, distribution sensitive welfare measures are infrequently used, not least because many of them have complex mathematical expressions or have units that are difficult for non-technical audiences to understand. While the concept of distribution sensitivity is well established and easily verified for a given welfare measure, the concept of *simplicity* is not. We propose a new axiomatic framework to formally define simplicity properties and study the constraints that these properties impose on welfare measures and the welfare orderings they represent. We use this framework to characterize a new simple distribution sensitive welfare measure that is subgroup decomposable: the "welfare gap", *the average factor by which individual incomes must be multiplied to attain a given reference income level*. We describe the welfare gap and its related inequality and poverty measures and present an empirical illustration using the global distribution of individual incomes from 1990 to 2019.

Keywords: Welfare Index, Poverty Index, Inequality Index, Distribution Sensitivity

JEL Classifications: D63, I32

---

* Corresponding Author: akraay@worldbank.org, 1818 H Street NW, Washington DC 20433.
1 World Bank; 2 University of Otāgo and Stanford University; 3 University of Oxford and University of Antwerp.
An earlier version of this paper was circulated as "A New Distribution Sensitive Index for Measuring Welfare, Inequality and Poverty", World Bank Policy Research Working Paper No. 10470. We thank Matthew Adler, Benu Bidani, Francois Bourguignon, Francisco Ferreira, Laurence Chandy, Gaurav Datt, Maya Eden, James Foster, Garance Genicot, Vickie Kraay, Peter Lanjouw, Max Lawson, Luis-Felipe Lopez-Calvo, Daniel Mahler, Roy Van der Weide, and seminar participants at several seminars, workshops, and conferences for helpful discussions and feedback. The names of the authors are in random order, using the author randomization tool of the American Economic Association, archived here. We acknowledge financial support from the UK government through the Data and Evidence for Tackling Extreme Poverty (DEEP) Research Program. The views expressed here are the authors', and do not necessarily reflect the official views of the World Bank, its Executive Directors, or the countries they represent. A verified reproducibility package for the working paper version of this paper is available here. Stata and R scripts for implementing the new welfare, poverty and inequality indices introduced in this paper are available here.

1 Introduction

Mean income per capita and the poverty headcount are among the most common measures used to summarize living standards and to assess progress towards national and international development goals. The popularity of these measures is largely due to their simplicity, which makes them easy to understand and communicate to wide audiences. However, these measures also have a fundamental limitation: they are silent on the extent of inequality within the population of interest, i.e., the entire population in the case of the mean, and the poor in the case of the headcount.

This limitation is well understood, and a long literature has generated an impressive array of distribution sensitive welfare and poverty measures to address it.[1] However, these distribution sensitive measures are much less frequently used in practice. To illustrate this, Table 1 displays the results of a simple keyword search for leading poverty and welfare measures in three bodies of documents: academic research papers in EconLit; policy documents of major multilateral development banks as curated at nlp4dev.org; and general writing in the Google N-gram Viewer. Across all three categories of documents, the frequency of references to the poverty headcount is two to three orders of magnitude larger than the frequency of references to the most common distribution sensitive poverty measures, the squared poverty gap and the Watts index. The pattern is similar for welfare measures in the bottom panel: references to average income are two orders of magnitude more common than references to the best-known distribution sensitive welfare measures, the Atkinson and Sen indices.

One likely reason for the limited uptake of distribution sensitive measures is that they often are difficult to explain in non-technical terms or lack intuitive units. For example, the mathematical formulation of the Atkinson index as a constant elasticity of substitution aggregate of individual incomes is likely to be daunting to most non-technical users – despite the elegant equally distributed equivalent income (EDEI) interpretation of this measure. Similarly, the Watts poverty index is measured in log-point differences from the poverty line that cannot easily be interpreted as percent differences unless they are small. For policymakers and the public, "average income in dollars" or "the number of poor people" are more intuitive, and therefore more likely to be effective in influencing policy and public opinion, despite their lack of distribution sensitivity. This dilemma is aptly summarized by the succinct remark in Watts (1969) that the headcount has "*little but its simplicity to recommend it.*"

[1] Early critiques of the headcount index for its lack of distribution sensitivity go back to Watts (1969) and Sen (1976). Distribution-sensitive poverty and welfare measures have been introduced by numerous scholars, including Watts (1969), Atkinson (1970), Sen (1976), Thon (1979), Takayama (1979), Kakwani (1980), Clark, Hemming, and Ulph (1981), Donaldson and Weymark (1980), Chakravarty (1983), Foster, Greer, and Thorbecke (1984), Hagenaars (1987), and Chakravarty (2009). We acknowledge that there is debate in the profession as to the need for such summary welfare measures, as opposed to relying on a "dashboard" of distinct measures of average income and income inequality. We do not enter these debates in this paper but rather take the need for a single summary distribution-sensitive welfare measure as given.

Table 1. Frequency of Distribution Sensitive Poverty and Welfare Measures in Academic, Policy, and Public Discourse

	Academic Use	Policy Use	Public Discourse
	*EconLit**	*Multilateral Development Banks at nlp4dev.org**	*Google N-gram***
Poverty Measures			
Poverty Rate OR Poverty Headcount ***	40832	26730	40.49
Squared Poverty Gap	473	681	0.32
Watts Index	47	52	0.09
Welfare Measures			
Average income	61493	15492	36.11
Atkinson Index	744	122	0.49
Sen Index	258	70	0.12

Notes: * Number of documents. ** 1 million times fraction of corpus in most recent year. *** Search results in last column for "poverty rate" only, as the Ngram viewer does not support Boolean search. Search results retrieved March 22-24, 2023. This table can be reproduced using the reproducibility package for the working paper version of this paper, downloadable at: https://reproducibility.worldbank.org/index.php/catalog/7.

In this paper we propose a new welfare measure that is both distribution sensitive and simple to understand. Distribution sensitivity is a well-established concept with standard definitions that can readily be verified given the mathematical expression for a welfare measure. Simplicity, on the other hand, is not. To make progress on this front, we develop a novel formalization of what it means for a welfare measure to be "simple". We distinguish between welfare *orderings* that rank income distributions, and welfare *measures* that provide a numerical representation of these orderings by mapping income distributions into scalar quantities, with larger (or smaller) values representing preferred distributions. These welfare measures may also depend on parameters with income units (such as a poverty line or another reference income level) as well as unitless parameters (such as a coefficient of inequality aversion). We rely on standard axioms such as Pigou-Dalton sensitivity to impose ethically desirable properties on welfare orderings.

We next propose a series of simplicity properties that impose restrictions on the welfare measures that represent these orderings – and indirectly restrict the orderings themselves. These simplicity properties capture key features of welfare measures that make it easier for users to understand and engage with them. Some simplicity properties restrict the units of the welfare measure to be intuitive (for example, the units of the welfare measure should be the same units as income, or the ratio of two incomes). Others restrict the sign and orientation of welfare measures (for example, the welfare measure should take on only positive values or should be increasing in incomes) or limit the complexity of the mathematical expression of the index (for example, by restricting the number of parameters). Yet other simplicity properties are intended to simplify analysis using the welfare measure (for example, the welfare measure should be subgroup decomposable with population weights, or welfare comparisons should not depend on certain parameters of the welfare measure, or the welfare measure should be invariant to the units in which income is measured).

We think this approach is useful for two reasons. First, the “simplicity” of any given welfare measure is, much like beauty, in the eye of the beholder. This makes it difficult adjudicate assertions that one welfare measure is “simpler” than another. Instead, we think it is more productive to narrowly define key attributes of simplicity by proposing well-defined simplicity properties. This allows us to assess the simplicity of a measure in terms of the set of simplicity properties that it satisfies. Second, breaking down simplicity into a series of properties recognizes that not all aspects of simplicity may be equally desirable, and allows different users to place different weights on various aspects of simplicity. For example, the orientation of a welfare measure may be essential for some users, while the units of the measure may be essential for others. This, in turn, implies that a measure that is “simple” for some users may not be “simple” for others, unless they agree on the set of simplicity properties.

We next characterize the set of welfare measures that satisfy various combinations of the simplicity properties we propose. Doing so exposes stark trade-offs between certain attributes of simplicity and ethical validity. Our first main finding is an impossibility result: a welfare measure that satisfies two basic simplicity properties, subgroup decomposability and income units, cannot represent a Pigou-Dalton-sensitive welfare ordering (Theorem 1). To escape this impossibility, we maintain the requirement of subgroup decomposability, but we relax the permissible units property to allow the welfare measure to either have income units or to be the ratio of two incomes. Our second main result provides a full characterization of welfare measures that satisfy these two simplicity properties and demonstrates that some of them represent Pigou-Dalton-sensitive welfare orderings (Theorem 2).

However, Theorem 2 also shows that even when welfare measures satisfy certain basic simplicity properties, they can still have rather daunting mathematical expressions that may be difficult to analyze or to communicate to wider audiences. To make progress, we consider the implications of introducing more simplicity properties. In Theorem 3, we show that the imposition of one additional simplicity property – that the measure has at most one parameter with income units – uniquely characterizes a set of welfare measures that are either linear functions of mean of individual incomes or inversely proportional to the harmonic mean of individual incomes. Among these, only the measure involving the harmonic mean also is distribution sensitive. This measure is the “welfare gap”.

The welfare gap is the average across all individuals of the ratio z/y_i, where y_i is the income of individual i and z is a reference income level measured in the same units as income (for example, dollars per day). Each term in the average is the factor by which the income of the corresponding individual needs to be multiplied to attain the reference income level, and the welfare gap is *the average factor by which individual incomes need to be multiplied to reach the reference income level*. It is straightforward to see that the welfare gap is subgroup decomposable with population weights; is measured in units that are the ratio of two incomes; depends on a single income parameter; and is Pigou-Dalton distribution sensitive. What is less obvious is that it is the *only* Pigou-Dalton-sensitive welfare measure that satisfies these three simplicity properties.

While the welfare gap is the unique Pigou-Dalton-sensitive welfare measure that satisfies these three simplicity properties, there are other combinations of simplicity properties that also uniquely characterize the welfare gap. We describe two other such combinations in our last main result (Theorem 4). This illustrates how users with different preferences over attributes of simplicity may nevertheless be able to agree on a common distribution sensitive welfare measure, in this case, the welfare gap. At the same time, the welfare gap of course does not satisfy *all* plausible simplicity properties. For example, a reasonable simplicity property is that welfare measures should increase when individual incomes increase. While the simple expedient of negating the prosperity gap would allow it to satisfy this property, this would violate yet another plausible simplicity property – namely that welfare measures should take on strictly positive values. This further illustrates the trade-offs between different attributes of simplicity.

The welfare gap depends on a single parameter with income units: the reference income level z. This parameter gives the welfare gap its interpretation as the factor by which incomes must be multiplied to attain this reference income level. Any value of z that is measured in the same units as income will serve this purpose and selecting higher or lower values simply scales the welfare gap up or down without affecting its ranking of income distributions. Although this rescaling of the index is substantively innocuous, the specific choice of the reference income level z can still play a useful role in enhancing the interpretability and communications impact of the welfare gap when it is set to some salient value. For example, the World Bank has recently adopted the welfare gap as its key institutional measure of "shared prosperity," setting the reference income level at $z = \$25/\text{day}$. The World Bank refers to this as the "*global average income shortfall from a prosperity standard of $25 per day,*" or the "*global prosperity gap*" for short.[2] In this context, the reference income level is salient because it roughly corresponds to the level of mean per capita household income at which countries cross into high-income status according to the World Bank's income classification. Naturally, in other applications, other values of the reference income level may be more salient.[3]

The welfare gap has several additional appealing features beyond the formal simplicity properties that it satisfies by construction. First, it admits a simple multiplicative decomposition into average income and a measure of inequality. The inequality measure is a variant of the welfare gap: the average factor by which incomes must be multiplied to reach mean income. This decomposition is helpful as it readily allows users to additively separate contributions of growth in mean income and changes in inequality to the movements of the welfare gap. Second, the inequality measure implied by the welfare gap admits an exact multiplicative between/within-group decomposition. Third, the welfare gap can also be transformed into a poverty index that satisfies the Sen (1976) focus axiom by interpreting z as a poverty line and restricting the measure to individuals below the poverty line. The resulting measure is closely related to the poverty measure introduced by Bosmans, Esposito and Lambert (2011).

The remainder of this paper proceeds as follows. Section 2 presents our formal simplicity properties and main theoretical results characterizing welfare measures satisfying various combinations of these properties. Readers uninterested in the formal axiomatic derivation of the welfare gap can skip directly to Section 3, where we summarize the welfare gap and its related inequality and poverty measures. In Section 4, we provide an empirical illustration of the welfare gap, documenting levels and trends calculated using the global interpersonal income distribution between 1990 and 2019, a period when income growth was strongly pro-poor. As a result, the welfare gap improved much more rapidly than world average income over this period – owing

[2] See https://scorecard.worldbank.org/en/scorecard/home.
[3] For example, Max Roser recently proposed an additional poverty line of $30 per day to measure global poverty (Our World in Data, 2021; The New York Times, 2024).

mainly to rapid growth in East and South Asia, with little progress in Sub-Saharan Africa. Section 5 offers brief concluding observations.

2. An Axiomatic Characterization of Simple Welfare Measures

In this section we provide an axiomatic characterization of simple welfare measures. We first state two standard axioms that place ethical restrictions on welfare orderings. We then describe and discuss the implications of various simplicity properties that impose restrictions on the welfare measures that represent these orderings. Throughout this section, we refer to some common existing welfare measures, as well as some hypothetical ones, to illustrate key properties. Appendix A provides a table summarizing these welfare measures and the ethical axioms and simplicity properties they satisfy.

2.1 Framework

Let $\aleph$ denote the set of natural numbers and let $R_{>0}$ denote the set of strictly positive reals. Let $Y = \cup_{N\in\aleph} (R_{>0})^N$ denote the set of all income distributions whose individual incomes are all strictly positive. Let $y = \left(y_1, \dots, y_{N(y)}\right)$ denote an income distribution across $N(y) \in \aleph$ individuals and let $\bar{y}$ denote mean income in distribution y.

Let O denote a complete, transitive, and continuous ordering on the set of income distributions Y. We denote by $y \succcurlyeq_O y'$ (resp. $y \succ_O y'$) the fact that y is weakly (resp. strictly) preferred to y' by ordering O. We constrain the set of admissible orderings by imposing two standard axioms, which encapsulate ethical views on how the ordering should rank distributions. This is the standard axiomatic approach to formalizing ethical validity in welfare economics (Fleurbaey and Maniquet, 2011; Piacquadio, 2017; Eden, 2020; Hufe et al, 2022). Let Ω denote the set of orderings that satisfy the following axiom:

- **Axiom 1: Monotonicity**. F*or all* $\delta > 0$ *and all* $y, y' \in Y$ *such that* $N(y) = N(y')$, $y'_i = y_i$ *for all* $i \neq i'$ *and* $y'_{i'} = y_{i'} + \delta$, *we have* $y' \succcurlyeq_O y$.

We define any ordering $O \in \Omega$ to be a "**welfare**" ordering.[4]

Let Ω^{DS} denote the subset of orderings in Ω that also satisfy the following axiom:

- **Axiom 2: Pigou-Dalton Transfer** F*or all* $\delta > 0$ *and all* $y, y' \in Y$ *such that* $N(y) = N(y')$, $y'_i = y_i$ *for all* $i \neq i', i''$ *and* $y'_{i'} = y_{i'} + \delta \leq y'_{i''} = y_{i''} - \delta$, *we have* $y' \succ_O y$.

We refer to any ordering $O \in \Omega^{DS}$ as a "**Pigou-Dalton-sensitive welfare ordering**". Clearly, a Pigou-Dalton-sensitive welfare ordering is a welfare ordering, but the converse is not necessarily true.

Our objective is to study the relationships and trade-offs that may exist between these ethical axioms and various aspects of simplicity. Unlike these axioms, the simplicity properties are not motivated by ethical considerations. To make this distinction clear, we do not impose these

[4] Observe that we merely impose a weak version of the Monotonicity axiom, which only requires a weak preference when one distribution is obtained from the other by strictly increasing the income of some individual. As a result, poverty measures also represent welfare orderings, which provides some benefits for our exposition. However, all our results would be unchanged if we instead imposed a strict Monotonicity axiom.

properties on orderings directly but on the measures that represent the orderings. This is important because simplicity mostly refers to attributes of the measure rather than attributes of the ordering itself. The simplicity properties are meant to encapsulate various attributes of welfare measures that make it easier to engage with them. However, as we shall see shortly, simplicity properties imposed on the measure that represents a welfare ordering can indirectly constrain the ethical properties of that ordering.

In our framework, a measure M is a parametric function defined on income distributions and may also depend on a list of normative income parameters, such as a poverty line or a reference income level. Let the parameter list be denoted by $z = \left(z_1, \dots, z_J\right) \in (R_{>0})^J$. The number J of parameters may depend on the measure considered, but for the sake of brevity, our notation ignores this potential dependence. In this section, when the measure admits a unique income parameter, we denote this parameter by z_1. A parametric measure M is a function $M: Y \times (R_{>0})^J \rightarrow R$. For two different $z, z' \in (R_{>0})^J$, the measure M may represent two different orderings on income distributions. For example, the comparison of poverty rates in two countries may be different when a different poverty line is used. We denote the measure obtained for a given value z for the income parameters by $M_z: Y \rightarrow R$ and the ordering that measure M_z represents by O_z. Two remarks on the parameters of welfare measures are in order.

First, we allow measures to also have unitless parameters, such as a coefficient that governs inequality aversion. We do not introduce notation for parameters without income units because our simplicity properties do not impose any direct constraints on them (although they can be constrained indirectly, as we discuss below).

Second, many mainstream measures admit income parameters. For example, poverty measures are based on a poverty line, while growth is computed from income in a reference period. Another example is the welfare index introduced in Watts (1969), which is the average of the logarithm of individual incomes expressed as a fraction of a reference income. Furthermore, many measures admit several income parameters. The most obvious cases (but not necessarily the most interesting) are indicators that aggregate several incomes to compute a summary income parameter. For instance, the \$25/day value of the reference income level used by the World Bank in its "prosperity gap" is the average across multiple countries of household income at the point where countries achieve high-income status. More interestingly, some indicators "irreducibly" depend on several income parameters, in the sense that they cannot be defined from a unique summary parameter that aggregates all income parameters. Poverty measures characterized by Decerf (2021), which depend on two poverty lines – one capturing subsistence and the other capturing social participation – are one example. Another example is provided by utilitarian social welfare functions (SWFs) based on piecewise linear utility functions, where any two consecutive linear segments are separated by some reference income level. These piecewise linear SWFs are ubiquitous in taxation, where they implicitly rationalize taxation schemes whose marginal rates differ across income brackets but are constant within brackets.

The simplicity properties we propose constrain both the functional form of a measure and, potentially, the acceptable values for its parameters. Let $P \subseteq (R_{>0})^J$ denote the subset of parametric values for which the parametric measure satisfies our axioms and the relevant simplicity properties.[5] Throughout, we require that if $z \in P$ then $kz \in P$ for all $k > 0$. For measures with a unique income parameter, i.e., $J = 1$, this innocuous assumption implies that $P = R_{>0}$.

We would like parametric measures to represent welfare orderings. It is useful for our purposes to distinguish two forms of representation: increasing measures have higher values for preferred distributions, while decreasing measures have lower ones for them. Mean income and the Atkinson index are examples of the former, while the poverty headcount is an example of the latter.

Definition 1: *M is a **parametric welfare measure** if either*

- *for all $z \in P$ there exists $O_z \in \Omega$ such that $y \succcurlyeq_{O_z} y' \iff M_z(y) \geq M_z(y')$ for all $y, y' \in Y$ (**increasing measure**), or*
- *for all $z \in P$ there exists $O_z \in \Omega$ such that $y \succcurlyeq_{O_z} y' \iff M_z(y) \leq M_z(y')$ for all $y, y' \in Y$ (**decreasing measure**).* [6]

*M is a **parametric Pigou-Dalton-sensitive welfare measure** if we replace Ω by Ω^{DS} in Definition 1.*

For the sake of terminological brevity, we will often simply refer to "measures" or "welfare measures" when it is not confusing to do so. Observe that one ordering can be represented by many different measures, and any two such measures must be monotonic transformations of each other. To discriminate between alternative measures, we formally define a set of properties encapsulating various attributes of simplicity. These properties determine which orderings can be represented by a measure that satisfies these properties. Thus, the theoretical results we present below highlight any trade-offs between conceptual validity (ethical axioms) of the underlying ordering and the usability (simplicity properties) of the associated measure that represents the ordering.

2.2 Two basic simplicity properties lead to an impossibility

Simplicity is notoriously difficult to formalize, which may explain the scarcity of this notion in axiomatic research on welfare indicators. One reason for this apparent paradox could be that simplicity is an attribute that is partly subjective, as it reflects individuals' backgrounds and their experiences with the measure. However, we argue that there are objective aspects of simplicity, which we seek to formalize via properties we propose below. Ideally, one would want to define simplicity properties that are sufficiently strong to meaningfully discriminate among measures but, at the same time, not so strong as to excessively restrict the set of orderings that can be represented by a simple measure. We also emphasize that we are not trying to propose a simple theory – nor an intricate one (the proofs are not particularly complex) – but one that convincingly formalizes simplicity and characterizes simple welfare measures.[7]

---

[5] The set P potentially depends on the combination of axioms and simplicity properties considered. See, for instance, Case 2 of Theorem 2 for an example, where set P is a strict subset of $(R_{>0})^J$.

[6] Implicitly, the definition of a welfare measure already assumes a simplicity property, namely that the direction of the measure cannot change with the value selected for its income parameters. In other words, a welfare measure M cannot be such that, for $z, z' \in P$, M_z is increasing while $M_{z'}$ is decreasing.

[7] It is important to distinguish between a simple property and a property that encapsulates simplicity. A property can be simple to understand, but at the same time unrelated to the simplicity of welfare measures that satisfy it. For

To monitor global poverty, Atkinson (2016) offers a synthesis of principles that should be followed in drawing up a portfolio of poverty indicators. These principles also highlight the tension between measures being simple yet satisfying generally accepted ethical axioms. For example, Atkinson's Principle 3 states that "The definition of the indicator should be generally accepted as valid and have a clear normative interpretation," while Principle 2 states that "The indicator should be transparent and identify the essence of the problem," adding that "it must have intuitive validity and be meaningful to the user."[8] Using our terminology, Atkinson (2016) proposes that indicators should satisfy both some basic ethical axioms (Principle 3) and some simplicity properties (Principle 2). In this sense, our framework to characterize "simple" welfare measures can be seen as an attempt at formalizing Atkinson's Principle 2.

The properties we propose capture distinct aspects of simplicity. One key aspect of simplicity relates to the communicability and understandability of the measure, i.e., how easy it is to describe its mathematical formula and provide meaning to the values obtained. Another is analytical simplicity, i.e., how easy it is to conduct analysis based on the measure. For example, analysis is simpler when the measure can be decomposed across subgroups or when its findings are robust to different values selected for its parameters. Relatedly, the measure is simpler to use when it is invariant to the income units used, say dollar a day or pesos a year. Some of the simplicity properties we propose are well-known, such as subgroup decomposability, while others are new, like those capturing intuitive units.

Our first property, Subgroup Decomposability, captures at least two aspects of simplicity. First, a measure that satisfies this property is easy to communicate because its mathematical formula must have the following form:

(1) $$M^{SubDec}(y,z) = \frac{1}{N(y)} \sum_{i=1}^{N(y)} f(y_i, z)$$

where function $f: R_{>0} \times (R_{>0})^J \rightarrow R$ provides the contribution of each individual and does not depend on the incomes of other individuals. Hence, any subgroup decomposable measure is simply the average of individual contributions to the measure. We believe this contributes to the wide use and communicability of subgroup decomposable poverty measures, like the Foster-Greer-Thorbecke class of indices (especially the poverty headcount). This property also simplifies analysis, as the welfare of the total population is the weighted average of the welfare of its subgroups. The analyst can thus easily compare welfare across gender, ethnicity, or regions and track overall progress as a function of progress in these subgroups, with each assigned their population weight.[9]

instance, a property requiring that $M(ky, kz) = M(y, z) + k$ is simple to understand, but it might not capture any relevant aspect of simplicity.

[8] In recommending the use of multidimensional poverty indicators based on the counting approach, the report also cites Alkire et al. (2015, p. 143) as stating "counting the number of observable deprivations in core indicators has an intuitive appeal and simplicity that has attracted not only academics but also policymakers and practitioners."

[9] An additional practical advantage of subgroup decomposability with population weights is that it simplifies the estimation of sampling error in the overall index.

- **SP1 – Subgroup Decomposability with Population Weights:** *For all $y \in Y$, all $z \in P$, and all population subgroups A such that $y = (y_A, y_{-A})$, we have*:

$$M(y,z) = \frac{N(y_A)}{N(y)} M(y_A, z) + \frac{N(y_{-A})}{N(y)} M(y_{-A}, z) \tag{2}$$

To be sure, *SP1* is a strong property. The Sen index and all Atkinson indices except mean income violate *SP1*. More generally, *SP1* prevents welfare orderings from depending on relative aspects of income, such as the Sen index, which depends on income ranks, assigning higher weights to individuals who occupy lower ranks in the income distribution. This provides a first illustration of our point that simplicity properties imposed on the measure that represents an ordering may constrain the ethical properties of the ordering itself. In this case, subgroup decomposability rules out ethical properties that posit that relative positions in an income distribution are welfare relevant, as argued for example by Fleurbaey, 2012.[10]

The second simplicity property is new. Its purpose is to constrain the units of the welfare measure. Welfare orderings are defined on a list of individual incomes, which all have common income units. It is natural and intuitive to require individual incomes to be aggregated in a way that preserves these units. From a communication perspective, the easiest case is when the measure itself has income units. In that case, the welfare level is easy to understand as it corresponds to some income level.

One simple example is the gap between individual incomes and some reference income level:

$$M^1(y,z) = \frac{1}{N(y)} \sum_{i=1}^{N(y)} (y_i - \bar{z}) = \bar{y} - \bar{z} \tag{3}$$

where income parameters z could be a reference income distribution, for example the income distribution in the United States, $z = y^{US}$, and thus $\bar{z}$ is mean income in the US. Many other existing welfare indices, such as the Atkinson indices, have income units. This is also the case for the following extension of Atkinson welfare indices that admits income parameters:

$$M^2(y,z) = A_\varepsilon(y,z) = \left(\frac{1}{N(y)} \sum_{i=1}^{N(y)} (y_i + \bar{z})^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}} \tag{4}$$

where ε is a unitless parameter governing aversion to inequality and M^2 corresponds to classical Atkinson indices when $\bar{z} = 0$. This provides an example of a welfare measure where the value of the income parameter affects how the measure orders distributions.

The Sen Index also has income units:

[10] In our framework, subgroup decomposability is not motivated from an ethical validity perspective, i.e., as a separability requirement, but for the aspects of simplicity that it encapsulates.

(5) $$M^3(y) = Sen(y) = \frac{2}{N^2(y) + N(y)} \sum_{i=1}^{N(y)} (N(y) + 1 - i) \times y_i = \bar{y} \times (1 - Gini(y))$$

where $Gini(y)$ denotes the well-known inequality measure and incomes are ordered such that $y_1 \leq y_2, \dots, \leq y_{N(y)}$.

However, the following welfare measure does not have income units:

(6) $$M^4(y) = \frac{1}{N(y)} \sum_{i=1}^{N(y)} {y_i}^{1/3},$$

because the units of M^4 are income units to the power one-third.

To formally distinguish between these possibilities, we must mathematically define what it means for a measure to "have income units." Axiomatic reasoning, which is conducted on the vector of real numbers, ignores the fact that some of these real numbers have units with meaning, such as incomes measured in dollars a day. The functional form of the measure can preserve (or dispense with) this meaning. Introducing this requirement necessitates a mathematical definition of "preservation." One cumbersome solution would be to formally define all the units' operations rules for all possible mathematical operations but, as we will see shortly, even this does not ensure that a measure has income units in a meaningful sense.[11] This illustrates how simplicity can resist formalization. Fortunately, for our purposes, we can propose a shortcut, which combines weakened versions of two properties found in the literature – e.g., in Foster and Szekely (2008) – to adequately capture what it means to have income units.

The first property is that the measure is homogeneous of degree one in all its arguments that have income units. This "homogeneity of degree one in income units" property implies that comparisons across distributions are unchanged when incomes are expressed in different income units, say dollars per day or pesos per year.

The second property restricts the values taken by the measure when all individual incomes are equal. One option would be to require this value to be equal to the common income level. In our terminology, the ensuing measures have "normalized income units." Note that such a normalization property would be quite strong as it rules out many measures that have income units. For instance, it rules out measures like M^1 and M^2. Although this option will be helpful in Theorem 4 for the definition of measures that have "normalized income units," we begin with a weaker normalization restriction: when the common income level increases from $a > 0$ to $b > a$, the measure M either increases or decreases by $b - a$, depending on whether the measure in question is an increasing measure, like M^1, or a decreasing measure, like the average poverty gap.[12]

With these ingredients, we can define whether a measure has income units. Let $\mathbb{1}_N$ denote a vector of N ones, so that for any $a > 0$ *we have* $a\mathbb{1}_N = (a, \dots, a)$.

---

[11] This would require formalizing mathematically acceptable additions, subtractions, multiplications, and divisions of quantities; dimensional consistency; conversion of units, etc.

[12] Of course, the average poverty gap violates this weaker restriction above the poverty line because it is unaffected by increases in individual incomes that take place above the poverty line.

Definition 2: *The parametric measure* M ***has income units*** *if*

- *for all* $k > 0$*, all* $y \in Y$*, and all* $z \in P$*, we have* $M(ky, kz) = kM(y, z)$ ***(homogeneity of degree one in income units****), and*
- *for all* $a, b > 0$ *with* $b > a$*, all* $N \in \aleph$*, and all* $z \in P$*, we have* $|M(b\mathbb{1}_N, z) - M(a\mathbb{1}_N, z)| = b - a$ *(****weak normalization****).*

The parametric measure M ***has normalized income units*** *if it has income units and*

- *for all* $a > 0$*, all* $N \in \aleph$*, and all* $z \in P$*, we have* $M(a\mathbb{1}_N, z) = a$ *(****strong normalization****).*

In Definition 2, the purpose of the absolute value function, denoted by $|.|$, is to allow for both increasing and decreasing welfare measures. Importantly, our characterization of the welfare gap would still hold if the absolute value function were removed, limiting this second property to increasing measures.

Definition 2 rules out measures with units that could be considered income units (under standard operations rules for units), but only in an "artificial" way that would make them hard to understand. Consider for instance,

$$M^5(y, z) = (\bar{z})^{2/3} \times \frac{1}{N}\sum_{i=1}^{N}(y_i)^{1/3} = (\bar{z})^{2/3} \times M^4(y) \tag{7}$$

where $\bar{z}$ is defined in the same way as in M^1. The units of M^5 can be seen as income units, in the sense that z and individual incomes all have income units and the "exponentiation" unit operation rule implies that $income\ units = (income\ units)^{2/3} \times (income\ units)^{1/3}$. However, M^5 clearly aggregates individual incomes in a way that does not respect their units, i.e., via M^4, and its final income units are artificially obtained using multiplication by a constant with appropriate (but unusual) units, namely $(\bar{z})^{2/3}$. Although M^5 satisfies homogeneity of degree one, it violates weak normalization and therefore does not have income units according to our Definition 2.

Our second simplicity property, which we call Strong Intuitive Units, is based on Definition 2 and requires the measure to have income units. It captures the idea that a measure with intuitive units is easier to communicate to policymakers and the public. Such communication and general acceptance are likely to be significantly impaired for a measure whose units might be, for example, a power of income units, such as the square root of dollars per day. Such measures are likely to be meaningful only to specialists, who are familiar with complex mathematical formulas and their purposes.

- **SP2 – Strong Intuitive Units:** *M has income units.*

Observe that utilitarian social welfare functions, i.e., $U(y) = \frac{1}{N}\sum_i u(y_i)$, based on a strictly concave utility function $u(y_i)$ violate *SP2*. The fact that they do not have intuitive units might be a reason why such social welfare functions are seldom used outside of academic circles.

We now show that these two simplicity properties, subgroup decomposability (*SP1*) and strong intuitive units (*SP2*), uniquely characterize a family of welfare measures, formalized in the following result:

Theorem 1: *M is a* ***parametric welfare measure*** *that satisfies SP1 and SP2 if and only if either $M = M^+$ or $M = -M^+$ for some increasing measure M^+ such that for all $y \in Y$ and all $z \in P$*

$$M^+(y,z) = \frac{1}{N}\sum_{i=1}^{N} y_i - B(z) = \bar{y} - B(z) \quad (8)$$

where the function $B: P \to R$ is homogeneous of degree one.

Proof: *see Online Appendix A*.

Theorem 1 essentially states that welfare measures satisfying *SP1* and *SP2* can only be linear translations of mean income.[13] This implies the following corollary:

Corollary 1: *No parametric Pigou-Dalton-sensitive welfare measure satisfies both SP1 and SP2.*

Proof: Directly follows from Theorem 1, as mean income is not Pigou-Dalton sensitive.

Theorem 1 and Corollary 1 illustrate a sharp trade-off between simplicity and ethical validity: it is impossible to have a Pigou-Dalton-sensitive welfare measure that simultaneously satisfies two basic simplicity properties, *SP1* (subgroup decomposability) and *SP2* (strong intuitive units). This is because the only welfare measure satisfying these properties is a linear translation of mean income, which is not Pigou-Dalton sensitive. This, in turn, tells us that our first two simplicity properties, while eminently reasonable on their own, are very restrictive when imposed together because they rule out distribution sensitive measures. Our goal in the next subsection is to escape this impossibility by proposing a slightly less demanding simplicity property to accommodate Pigou-Dalton-sensitive welfare measures.

2.3 Escaping the impossibility by weakening Strong Intuitive Units

To escape the impossibility stated in Corollary 1, we weaken the requirement that the measure has income units (*SP2*). In Section 2.6, we consider the implications of taking an alternative route, namely keeping *SP2* and relaxing subgroup decomposability (*SP1*).

We relax *SP2* because we believe it is possible to do so without losing too much of the simplicity encapsulated in it. Specifically, we expand the permissible units for the welfare measure to include measures that are the ratio of two incomes. For instance, such a measure could rescale mean income by mean income in a reference distribution, say that of the US, e.g., $\frac{\bar{y}^{US}}{\bar{y}}$ or $\frac{\bar{y}}{\bar{y}^{US}}$. We believe that such units remain intuitive because it is easy to understand that mean income in the US is twice or half as large as mean income in the distribution considered.

Our weakening of *SP2* relies on the following definition. This definition states that a measure M is a "ratio of two incomes" when it is defined as a ratio based on an underlying measure M^* that has income units computed for the income distribution considered (y) and for a reference income distribution, denoted by $z^{ref} \in Y$. The reference income distribution z^{ref} belongs to the income parameters of measure M, but does not belong to the income parameters of measure M^*.

[13] The functional form of $B(z)$ is of secondary importance because it merely defines a constant for a given value selected for the income parameters.

Definition 3: *The parametric measure* M ***is the ratio of two incomes*** *if for some partition of its income parameters* $z = (z^{ref}, z^*)$ *there exists a parametric measure* M^* *that has income units such that either:*

(9)
$$M(y,z) = \frac{M^*(y,z^*)}{M^*(z^{ref},z^*)} \qquad for\ all\ y \in Y\ and\ all\ z \in P \qquad or$$
$$M(y,z) = \frac{M^*(z^{ref},z^*)}{M^*(y,z^*)} \qquad for\ all\ y \in Y\ and\ all\ z \in P.$$

The parametric measure M ***is the ratio of two normalized incomes*** *if in addition measure* M^* *has normalized income units.*

Observe that both the numerator and the denominator are based on the same measure M^* that has income units, but one is computed for the income distribution under consideration while the other is computed for a reference income distribution, namely $z^{ref} \in Y$. For instance, we could have $z^{ref} = y^{US}$ and $M^*(y,z^*) = \bar{y}$, which results in $M(y,z) = M(y,z^{ref}) = M(y,y^{US}) = \frac{\bar{y}}{\bar{y}^{US}}$ or $\frac{\bar{y}^{US}}{\bar{y}}$ as per our earlier example. The reference distribution z^{ref} is kept fixed when the measure M compares any two income distributions $y, y' \in Y$. Note also that this definition is stronger than simply requiring M to have units that are a ratio. For example, the poverty headcount is a ratio, but it is a ratio of the number of people below the poverty line to the total number of people, while Definition 3 requires the welfare measure to be a ratio of incomes because M^* has income units.

Our next simplicity property, which we call Weak Intuitive Units, extends the set of allowable units to include not just income units as in *SP2*, but also the ratio of two incomes:

- **SP3 – Weak Intuitive Units:** *M has either income units or is the ratio of two incomes.*

SP3 increases the set of permissible units for welfare measures, meaning that it is "weaker" (i.e., less restrictive) than *SP2* in an axiomatic sense, while still retaining much of its simplicity. Arguably, if welfare measures with income units are simple to understand, then so are welfare measures that are the ratio of two incomes.

Theorem 2 fully characterizes the ensuing welfare measures.

Theorem 2: *M is a parametric* ***welfare measure*** *that satisfies SP1 and SP3 if and only if:*

- *$M = M^+$ or $M = -M^+$ for some increasing measure M^+ such that for all $y \in Y$ and all $z \in P$*

(10)
$$M^+(y,z) = \frac{1}{N}\sum_{i=1}^{N} y_i - B(z) = \bar{y} - B(z)$$

where the function $B: P \to R$ is homogeneous of degree one ***(Case 1); or***

- *for some partition of income parameters $z = (z^{ref}, z^*)$ and for all $y \in Y$ and all $z \in P$,*

$$M(y,z) = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)} = \frac{Mean\big(y - B(z^*)\mathbb{1}_{N(y)}\big)}{Mean\big(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\big)} \quad (11)$$

where the function $B: P \to R$ is homogeneous of degree one and either $\bar{z}^{ref} > B(z^*)\ \forall z \in P$ (M is an increasing measure) or $\bar{z}^{ref} < B(z^*)\ \forall z \in P$ (M is a decreasing measure) ***(Case 2)****; **or***

- *for some partition of income parameters $z = (z^{ref}, z^*)$ and for all $y \in Y$ and all $z \in P$*

$$M(y,z) = \frac{\frac{1}{N(y)}\sum_i \frac{1}{y_i - B(z^*)}}{\frac{1}{N(z^{ref})}\sum_j \frac{1}{z_j^{ref} - B(z^*)}} = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\big(y - B(z^*)\mathbb{1}_{N(y)}\big)} \quad (12)$$

where the function $B: P \to R_{\leq 0}$ is homogeneous of degree one and $Harm(y)$ denotes the harmonic mean ***(Case 3)****.*

Proof: *see Online Appendix A*.

Case 1 in Theorem 2 relates to welfare measures that have income units. In turn, ***Cases 2*** and ***3*** relate to welfare measures that are the ratio of two incomes. More importantly, Theorem 2 shows that there are essentially two types of welfare measures that satisfy both SP1 and SP3. The first type consists of linear transformations of mean income (***Cases 1*** and ***2***). The second type consists of measures that are proportional to the inverse harmonic mean of linearly translated incomes (***Case 3***).

Theorem 2 calls for two remarks. First, weakening *SP2* into *SP3* allows us to escape the impossibility derived in Theorem 1 and Corollary 1. That is, while there are no Pigou-Dalton-sensitive welfare measures that satisfy both *SP1* and *SP2*, there *are* Pigou-Dalton sensitive measures that satisfy *SP1* and *SP3*. These are the measures in Case 3, since the harmonic mean is Pigou-Dalton sensitive. Second, a quick glance at the expressions in Theorem 2 suggests that not all these welfare measures would qualify as easy to understand or communicate. This is particularly the case for some of the measures corresponding to Case 3. This shows that *SP1* and *SP3* may not be sufficient to characterize a Pigou-Dalton-sensitive welfare measure that also is easy to communicate and engage with. Additional simplicity properties are needed to further discriminate across measures.

We do this by introducing one additional simplicity property that requires the measure to have at most one parameter with income units. From the perspective of the policymaker, this property simplifies the implementation of the measure as a value must be selected for only one parameter. This requirement also makes the measure easier to communicate since only one such parameter needs to be described. As before, this property does not prevent the measure from having additional parameters that are unitless.

- **SP4 – At Most One Income Parameter:** *M admits at most one income parameter z_1, i.e., $J \in \{0,1\}$.*

Theorem 3 shows that imposing *SP4* considerably simplifies the mathematical expressions for the resulting welfare measures.

Theorem 3: M is a parametric **welfare measure** *that satisfies SP1, SP3 and SP4 if and only if*

- $M = M^+$ *or* $M = -M^+$ for some welfare measure M^+ such that for all $y \in Y$ and all $z_1 > 0$ we have $M^+(y, z_1) = \bar{y} - B(z_1)$ where function $B: P \rightarrow R$ is *homogeneous of degree one* ***(Case 1)****;* ***or***
- $M(y, z_1) = \frac{\bar{y}}{z_1}$ *for all* $y \in Y$ *and all* $z_1 > 0$ ***(Case 2)****;* ***or***
- $M(y, z_1) = \frac{z_1}{Harm(y)}$ *for all* $y \in Y$ *and all* $z_1 > 0$ ***(Case 3)****.*

Proof: *see Online Appendix A.*

Theorem 3 calls for two remarks. First, the formulas for the corresponding welfare measures appear considerably easier to communicate than those obtained in Theorem 2, as the welfare measures are either a translation of or proportional to mean of individual incomes, or inversely proportional to their harmonic mean. Second, while the welfare measures in Cases 1 and 2 are not distribution sensitive, the welfare measure corresponding to Case 3 is Pigou-Dalton sensitive. This welfare measure corresponds to the "welfare gap", defined as follows:

Definition 4: The parametric measure $W(y, z_1)$ *is the* ***welfare gap*** *if* $W(y, z_1) \equiv \frac{1}{N}\sum_{i=1}^{N}\frac{z_1}{y_i}$ *for all* $z_1 > 0$.

From Theorem 3, it is straightforward to see that the welfare gap is the only Pigou-Dalton sensitive welfare measure satisfying three key simplicity properties SP1, SP3 and SP4. We state this implication formally as the following corollary to Theorem 3:

Corollary 2: M *is a parametric Pigou-Dalton-sensitive welfare measure that satisfies SP1, SP3 and SP4 if and only if* M *is the welfare gap.*

Proof: Directly follows from Case 3 of Theorem 3 (as mean income is not Pigou-Dalton sensitive) and Definition 4.

To the best of our knowledge, the welfare gap is a new *measure*, although the *ordering* it represents is not new. Specifically, the welfare gap represents an Atkinson ordering with a degree of inequality aversion equal to two, because it is a monotonic transformation of an Atkinson welfare measure with this parameter restriction, i.e., $W(y, z_1) = \frac{z_1}{A_2(y)}$, where $A_\varepsilon(y) = \left(\frac{1}{N}\sum_{i=1}^{N} y_i^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}$.

As we noted earlier, many measures can represent the same welfare ordering. For example, Atkinson (1970) notes that the Atkinson ordering also is represented by the following isoelastic utilitarian social welfare function:

$U_\varepsilon(y) = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i^{1-\varepsilon}-1}{1-\varepsilon}$. Specializing to the case of $\varepsilon = 2$, the welfare gap is a monotonic transformation of this measure, i.e., $W(y, z_1) = -z_1(U_2(y) - 1)$.[14]

However, the Atkinson welfare measure satisfies *SP3* and *SP4*, but not subgroup decomposability (*SP1*), while the isoelastic utilitarian social welfare function satisfies *SP1* and *SP4,* but not weak intuitive units (*SP3*). In contrast, the welfare gap uniquely satisfies all three simplicity properties, *SP1*, *SP3*, and *SP4*. If lack of simplicity is an impediment to the diffusion of Pigou-Dalton-sensitive measures, there may be value in proposing a simpler – in the sense of satisfying these three properties – recalibration of existing measures, such as the welfare gap.

Theorem 3 also shows that we cannot freely select the degree of inequality aversion of a Pigou-Dalton sensitive measure that satisfies the combination of simplicity properties considered: the simplicity properties imply that the coefficient of inequality aversion of the Atkinson ordering must be equal to two. This again illustrates how simplicity properties imposed on welfare measures may indirectly restrict the ethical properties of the orderings they represent. However, we do not think this restriction is too costly, as this specific degree of inequality aversion does not appear to be implausible. For example, Kot and Paradowski (2022) provide direct empirical cross-country evidence and find a global mean estimate of inequality aversion of 1.92 (their Table 5). Similarly, Sterck (2024) finds a mean estimate of inequality aversion of 2.11 using survey data from experts and 2.41 using survey data from the general public in Kenya, India, South Africa, and the US.[15] If we interpret ε as a measure of risk aversion in the iso-elastic utilitarian welfare function $U_\varepsilon(y)$, we note that $\varepsilon = 2$ is appealing because it is widely used as a central value for risk aversion in calibrations of macroeconomic models.

2.4 Alternative characterizations of the welfare gap

In this section, we show that there are two different combinations of additional simplicity properties, which also fully characterize the welfare gap, thus providing alternative justifications for it.

One further simplicity property is that a welfare measure represents a welfare ordering that does not depend on the value selected for its income parameters. Being invariant to income parameters captures simplicity in the sense that it simplifies the definition of the measure as well as analysis with this measure. Indeed, the fact that poverty comparisons depend on the poverty line chosen complicates the task of the analyst, who must conduct robustness checks for different values of the poverty line. This also complicates the task of the policymaker, whose choice of the parameter is much more scrutinized and potentially contested. The long-lasting debates related to the choice of

[14] We note that the welfare gap has a superficial resemblance to the index proposed by Watts (1969), $WI(y, z_1) = \frac{1}{N}\sum_{i=1}^{N} ln(y_i/z_1)$, in that it the welfare gap depends on the ratio z_1/y_i while the Watts index depends on the ratio y_i/z_1. However, the Watts index represents a different welfare ordering, namely an Atkinson ordering with a coefficient of inequality aversion equal to one. Like the welfare gap, the Watts index satisfies subgroup decomposability (*SP1*). However, unlike the welfare gap, $WI(y, z)$ violates weak intuitive units (*SP3*), and its units of log differences from the income standard can only be interpreted as percent changes if they are small.

[15] Eden and Freitas (2024a, 2024b) use consumption patterns to calibrate an SWF for the U.S. and separately for a sample of low- and middle-income countries (LMICs). They find that the SWF is well-approximated by an Atkinson Index with a somewhat lower inequality aversion parameter between 0.6 and 1.2 (1.6) for the U.S. (LMICs).

poverty lines, both at national and global levels, illustrate the complexity of selecting an income parameter that affects normative comparisons.[16]

Yet, we might want some parameters to affect welfare comparisons, perhaps most importantly the parameter that tunes inequality aversion. Crucially, this is not precluded by the next property. The only constraint is that, if a parameter influences comparisons, then this parameter is unitless, like the coefficient of inequality aversion.

- **SP5 – Invariance to Income Parameter Values:** F*or all* $y, y' \in Y$ *and all* $z, z' \in P$*, we have*

$$M(y, z) \geq M(y', z) \Leftrightarrow M(y, z') \geq M(y', z')$$

Theorem 1 shows that welfare measures that satisfy *SP1* and *SP2* automatically satisfy *SP5*. Theorem 3 reveals that this implication also extends to welfare measures that satisfy SP1, SP3 and *SP4*: indeed, all measures characterized in Theorem 3 are monotonic transformations of either the mean or the harmonic mean. Fundamentally, the reason for the extension of this implication is that the unique income parameter of measures that are the ratio of two incomes must be z^{ref} rather than z^* (see Definition 4).[17] If z^{ref} was undefined, then these measures would also be undefined and could not represent a welfare ordering.

Although *SP5* is defensible from a simplicity perspective, it is undeniably a strong axiom from an ethical perspective. For instance, *SP5* prevents welfare comparisons from depending on a subsistence income level, below which incomes are treated differently than incomes above it. Thus, it precludes poverty measures. However, note that poverty measures are already excluded by some of the most basic requirements introduced earlier: most prominently, poverty measures do not satisfy *SP3* (nor *SP2*) because they are unaffected by changes in incomes taking place above the poverty line. For the same reason, they do not qualify as Pigou-Dalton sensitive welfare measures, and they violate the strong version of the monotonicity axiom (*Axiom 1*, where $\succcurlyeq_O$ is replaced by $\succ_O$). This is particularly the case for the headcount ratio, which is not even monotonic below the poverty line.

A testament to the fact that *SP5* embodies simplicity is that most mainstream welfare indices, like those in the Atkinson class or the Sen index, satisfy it. Indeed, these indices do not even rely on an income parameter and thus trivially satisfy *SP5*. One advantage of our attempt to conceptualize simplicity is to make such properties explicit. Among the illustrative measures considered above, M^2 violates *SP5*, but M^1, M^3, and M^4 satisfy it.

Our next result provides two alternative characterizations of the welfare gap. The second of these two characterizations is based on a narrower definition of measures that are the ratio of two incomes, namely on measures that are the ratio of two *normalized* incomes (Definition 3). The following simplicity property, which is logically stronger (i.e., more restrictive) than *SP3*, captures this.

[16] See Decerf (2025) on the longstanding debates around the definition of a global poverty line.

[17] Observe that we have $M^*(z^{ref}) = z^{ref}$ for any single parameter $z^{ref} > 0$ because M^* has income units and thus M^* is homogeneous of degree one in incomes and satisfies weak normalization.

- **SP6 – Normalized Intuitive Units:** M has either normalized income units or is the ratio of two normalized incomes.

Theorem 4 offers two alternative combinations of simplicity properties that characterize the welfare gap.

Theorem 4: *Let M be a parametric welfare measure. The three following statements are equivalent:*

1. *M is Pigou-Dalton sensitive and satisfies SP1, SP3 and SP5.*
2. *M is Pigou-Dalton sensitive and satisfies SP1 and SP6.*
3. *For some partition of its income parameters $z = (z^{ref}, z^*)$, where possibly $z^{ref} = z$, we have for all $z \in P$ that $M(y, z) = W(y, z_1)$ where $z_1 = Harm(z^{ref})$.*

Proof: *See Online Appendix A*.

The last statement in Theorem 4 says that the welfare measure is always equal to the welfare gap and specifies how its income parameters combine into the unique income parameter of the welfare gap. The first two statements in Theorem 4 provide alternative combinations of simplicity properties to Theorem 3 that imply the welfare gap.

All three characterizations of the welfare gap offered in Theorems 3 and 4 rely on subgroup decomposability (*SP1*) and on some form of restrictions on the units of the welfare measure – either weak intuitive units (*SP3*) or normalized intuitive units (*SP6*). The characterizations in Theorem 3 and the first statement of Theorem 4 rely on restrictions on how the welfare measure depends on income parameters, either restricting the measure to have only one income parameter (*SP4*) or insisting that the welfare measure is ordinally invariant to the income parameters (*SP5*). The second characterization in Theorem 4 dispenses with restrictions on the income parameters but instead imposes a stronger restriction on units than the first two (*SP6*).

2.5 Additional simplicity properties

Clearly, the welfare gap does not satisfy all the simplicity properties one could think of. For instance, the welfare gap violates a simplicity property that could also be considered relevant when measuring welfare, because it is a decreasing welfare measure. This might raise its communication costs, as it makes it necessary to mention that the lower values of the index correspond to higher values of welfare. The next property formalizes this aspect of simplicity. It requires that the direction of the measure is congruent with the objects being compared. Income is a "good," so any increase in income is considered an improvement. The welfare measure has congruent direction if its value increases when income increases, that is, when it is an increasing measure.

- **SP7 – Increasing measure:** *M is an increasing measure as described in Definition 1.*

As formalized by Corollary 3 below, which directly follows from Theorem 2 above, there exists no *increasing* welfare measure that is subgroup decomposable (*SP1*), has weak intuitive units (*SP3*), and is Pigou-Dalton sensitive:

Corollary 3: *No* parametric *Pigou-Dalton-sensitive* **welfare measure** *satisfies SP1, SP3 and SP7.*

Proof Corollary 3: Follows immediately from Theorem 2.

Admittedly, there is an increasing Pigou-Dalton sensitive welfare measure that satisfies *SP1* and "almost" satisfies *SP3*: i.e., $-W = -\frac{z_1}{Harm(y)}$, is increasing and Pigou-Dalton sensitive. This welfare index corresponds to the welfare gap "up to a sign." Yet, this measure violates another possibly desirable simplicity property, which requires that the measure does not take negative values, encapsulated by the following simplicity property.

- **SP8 – Non-negative values:** M does not take negative values, i.e., $M: Y \times (R_{>0})^J \to R_{\geq 0}$.

This illustrates further the trade-offs between simplicity properties, as well as those between simplicity and ethical validity. Mean income is increasing and non-negative, but it is not Pigou-Dalton sensitive. In turn, the welfare gap is non-negative and Pigou-Dalton sensitive, but it is decreasing. Finally, the welfare gap multiplied by minus one is increasing and Pigou-Dalton sensitive, but it yields negative values.

The welfare gap is decreasing, which is a limitation to its simplicity, but we observe that non-congruent indicators are common in other contexts. For example, life expectancy at birth is a well-known measure of mortality. Mortality, as observed by the death of individuals, is a "bad", yet life expectancy at birth is an increasing measure. The success of this non-congruent measure might be partly explained by the fact that it admits a simple interpretation, namely the average lifespan of a fictitious cohort of newborns facing current age-specific mortality levels.

We close this section with one additional simplicity property satisfied by the welfare gap. Because the welfare gap is the ratio of two incomes, it does not depend on the income units in which incomes are expressed. In other words, their value is unchanged when moving from dollars per day to pesos per year. This property simplifies the dissemination of the measure to audiences using different income units.

- **SP9 – Invariance to Income Units:** F*or all* $k > 0$*, all* $y \in Y$*, and all* $z \in P$*,* $M(ky, kz) = M(y, z)$

SP9 imposes homogeneity of degree zero in income units. Thus, *SP9* rules out welfare measures that have income units (***Case 1*** in Theorem 2), which must be homogeneous of degree one in income units (Definition 2). This once again illustrates tradeoffs between simplicity properties.

2.6 Escaping the impossibility while keeping Strong Intuitive Units

Instead of relaxing the strong intuitive units property (*SP2)*, one might consider escaping the impossibility encapsulated in Corollary 1 by relaxing subgroup decomposability (*SP1)*. For example, we could replace *SP1* with a weaker ordinal decomposability property such as Subgroup Consistency. As shown by Foster and Szekely (2008), all Atkinson measures satisfy Subgroup Consistency. As they also satisfy *SP2* and are Pigou-Dalton sensitive for $\varepsilon > 0$, they offer an alternative to the welfare gap in escaping the impossibility encapsulated in Corollary 1. They also have the added benefit of being increasing measures (*SP7)*.

Nevertheless, we believe that insisting on subgroup decomposability (*SP1*) and relaxing strong intuitive units (*SP2*) – which leads us to the welfare gap – is preferable, for the following reason. An attractive feature of the welfare gap is that it is simple to describe its mathematical formula in words. The welfare gap inherits this feature from the *combination* of the two main simplicity properties. First, as explained above, any subgroup decomposable measure (*SP1*) can be described as the average of individual contributions to the measure because $M^{SP1}(y,z) = \frac{1}{N(y)} \sum_{i=1}^{N} f(y_i, z)$ where $f(y_i, z)$ denotes the individual contribution. Second, the weak intuitive units property (*SP3*) implies that the individual contributions must be defined either as incomes or as ratios of two incomes. Thus, the *combination* of SP1 and SP3 makes the welfare gap easy to describe as the average factor by which individual incomes need to be multiplied to attain the reference income, z_1. In contrast, a subgroup decomposable measure where individual contributions lack intuitive units is likely to be difficult to describe to a lay audience. For example, the units of the subgroup decomposable measure M^4 are income units to the power one-third, which is not particularly intuitive.

3 The Welfare Gap, Inequality, and Poverty

In this section, we discuss inequality and poverty measures that are related to the welfare gap. In Section 3.1, we begin with a brief recap of the welfare gap for the benefit of readers who may have skipped its formal axiomatic characterization in Section 2. Section 3.2 shows how the welfare gap can be decomposed into the product of mean income and an inequality measure, which in turn admits a within-between group decomposition. Section 3.3 shows how the welfare gap can be transformed into a poverty measure that satisfies the conventional focus axiom.

3.1 The Welfare Gap

The welfare gap is given by:

$$W(y,z) = \frac{1}{N} \sum_{i=1}^{N} \frac{z}{y_i} \tag{13}$$

where $i = 1, \dots, N$ indexes individuals; $y_i > 0$ denotes individual i's income; $y \equiv (y_1, \dots, y_N)$ denotes the distribution of income; and $z > 0$ is a fixed scalar reference income level, measured in the same units as y_i, that can be set to any salient value.[18]

The ratio z/y_i is the factor by which the income of person i must be multiplied to attain the reference income level. This ratio is decreasing in the income of person i, so the smaller this ratio, the higher is welfare. The value of this ratio tends to zero as the income of person i tends to infinity. The welfare gap is the average of these ratios across individuals, i.e., it is *the average factor by which individual incomes need to be multiplied to reach the reference level of income*. Because it is decreasing in incomes, higher values of the $W(y,z)$ correspond to lower levels of welfare. The "gap" in welfare gap emphasizes that reductions in $W(y,z)$ correspond to improvements in welfare.

Since the ratio z/y_i is a convex function of income, the welfare gap is distribution sensitive. We have already seen in Section 2 that the welfare gap is *Pigou-Dalton sensitive*. It is straightforward to verify that it also satisfies *transfer sensitivity*, recording improvements in welfare from

[18] For notational convenience in the rest of this paper, we refer to this unique income parameter as z rather than z_1, as in the notation of Definition 4. Reference income levels or reference points are common in the inequality literature (see e.g., Cowell and Ebert 2004; Cowell and Flachaire 2021) but less frequent in the literature on welfare measurement (Watts 1969 is an exception).

progressive transfers that occur lower in the income distribution (Kakwani, 1980), and *growth sensitivity*, recording improvements in welfare from growth that occurs lower in the income distribution (Fleurbaey and Michel, 2001; Dollar et al., 2015; Ray ⓡ Genicot, 2022).

The welfare gap depends on a reference income level z that provides it with readily interpretable units. The units of $1/y_i$ are "inverse currency units" which are not very intuitive, but z/y_i is a straightforward ratio reflecting the factor by which incomes must be multiplied to reach the reference income level. Changing the value of z to z' simply rescales the welfare gap by a factor of z'/z but does not affect how it ranks income distributions.

Since the welfare gap is a simple average of ratios, it is sub-group decomposable with population weights. This feature is of considerable value for policy applications because, for example, it means that the welfare gap at the global level is simply the population-weighted average of the same welfare gaps measured at the country level – a property we illustrate empirically in Section 4.

3.2 Contributions of mean income and inequality to the welfare gap

The welfare gap $W(y,z)$ admits a natural decomposition into average income and its dispersion:

$$W(y,z) = \left(\frac{z}{\bar{y}}\right) I(y,\bar{y}), \quad I(y,\bar{y}) \equiv \frac{1}{N}\sum_{i=1}^{N}\frac{\bar{y}}{y_i}, \tag{14}$$

The first term, $\frac{z}{\bar{y}}$, represents the shortfall of average income $\bar{y}$ from the reference income level z. The larger is mean income, the smaller is $\frac{z}{\bar{y}}$, and the smaller is the welfare gap $W(y,z)$.

The second term is a new and intuitive index of inequality measuring *the average factor by which individual incomes must be multiplied to reach mean income*, recognizing that incomes above the mean must be multiplied by a factor less than one. $I(y,\bar{y})$ is equal to one when incomes are distributed equally, i.e., when $y_i = \bar{y}$ for all individuals, with higher values of $I(y,\bar{y})$ corresponding to greater inequality. As inequality, measured by $I(y,\bar{y})$, declines, the welfare gap, $W(y,z)$ becomes smaller, i.e., welfare improves. We refer to $I(y,\bar{y})$ as the "mean ratio deviation," since it measures deviations from the mean as a ratio of the mean to individual incomes (while the mean log deviation measures them as log differences).

The mean ratio deviation has an intuitive interpretation in terms of equally distributed equivalent income (Atkinson, 1970). Specifically, $\bar{y}^* \equiv \frac{z}{W(y,z)}$ represents the level of income that, if earned by all individuals, generates the same level of welfare as $W(y,z)$, since $I(y^*,\bar{y}^*) = 1$ where $y^* \equiv (\bar{y}^*, \dots, \bar{y}^*)$ denotes the counterfactual distribution in which everyone earns $\bar{y}^*$. From Equation (14), $\frac{z}{W(y,z)} = \frac{\bar{y}}{I(y,\bar{y})}$. As a result, $I(y,\bar{y})$ can be interpreted as the factor by which mean income could be reduced if it were equally distributed across all individuals while still delivering the same level of welfare as $W(y,z)$.

The mean ratio deviation admits an exact multiplicative decomposition into within-group and between-group inequality. Formally, partition y into $g = 1, \dots, G$ mutually exclusive and exhaustive groups with N_g members. In addition, let y_g denote the distribution of income in group

g, and $\bar{y}_g$ the mean of this distribution. With this notation, the mean ratio deviation can be decomposed as follows:

$$(15)\qquad I(y,\bar{y}) = \underbrace{\left(\sum_{g=1}^{G} \frac{N_g}{N}\frac{\bar{y}}{\bar{y}_g}\right)}_{\substack{between-group\\ inequality\ I_{btw}}} \underbrace{\left(\sum_{g=1}^{G} w_g\, I(y_g,\bar{y}_g)\right)}_{\substack{within-group\\ inequality\ I_{wth}}}, \quad w_g = \frac{N_g/\bar{y}_g}{\sum_{g'=1}^{G} N_{g'}/\bar{y}_{g'}}$$

The first term captures between-group inequality as the population-weighted average of the gap between each group's mean income and overall mean income, with the gap measured as a ratio. Equivalently, this term corresponds to applying the mean ratio deviation to a counterfactual distribution in which inequality within groups has been eliminated by setting everyone's income equal to their group average income. The second term captures the within-group contribution to overall inequality as a weighted average of the mean ratio deviation within each group. The weights w_g sum to one and are greater for groups that are larger and/or have lower mean income.[19] If incomes are equally distributed within all groups, the second term is equal to one. Overall inequality is simply the product of these two terms.

Just as sub-group decomposability of the welfare gap $W(y,z)$ is an attractive property that makes it more useful for communications and policy purposes, the existence of an analytical between/within-group decomposition does the same for the mean ratio deviation $I(y,\bar{y})$. We highlight this feature of the mean ratio deviation in our empirical illustration in Section 4. In contrast, the most common inequality measures, like the Gini index and variants of the Kuznets ratio (of average income in the top X to the bottom Y percent of the income distribution), do not have this attractive property. Finally, while members of the Generalized Entropy class of inequality measures do admit a between/within-group decomposition, their units are considerably more difficult to interpret.[20]

3.3 Related Poverty Measure

The welfare gap is inclusive in the sense that it is affected by the incomes of all individuals. In this subsection, we briefly note that $W(y,z)$ can readily be transformed into a poverty measure $P(y,z)$ that satisfies the Sen (1976) focus axiom by interpreting z as a poverty line and restricting the index to individuals below the poverty line, as follows:

[19] The decomposition into within- and between-group inequality of many Generalized Entropy Class inequality measures also use weights that depend on the size and mean income of each group.

[20] The mean ratio deviation is monotonically related to three existing inequality measures. First, given the relationship between the welfare gap $W(y,z)$ and the Atkinson welfare measure, it is not surprising that it is related to the Atkinson inequality index $AI(y,\varepsilon) \equiv 1 - \frac{1}{\bar{y}}\left(\frac{1}{N}\sum_{i=1}^{N} y_i^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}$; specifically, $I(y,\bar{y}) = \left(1 - AI(y,2)\right)^{-1}$. Second, $I(y,\bar{y})$ is related to the Shorrocks (1980) Generalized Entropy class of inequality measures $GE(y,\alpha) \equiv \frac{1}{\alpha(\alpha-1)}\frac{1}{N}\sum_{i=1}^{N}\left(\left(\frac{y_i}{\bar{y}}\right)^{\alpha} - 1\right)$; specifically, $I(y,\bar{y}) = 2GE(y,-1) + 1$. Third, $I(y,\bar{y})$ is related to a proposal in Dalton (1920) to measure inequality as $D(y,c) \equiv \left(c - \left(\frac{1}{N}\sum_{i=1}^{N} y_i\right)^{-1}\right) / \left(c - \left(\frac{1}{N}\sum_{i=1}^{N} y_i^{-1}\right)\right)$ where c is a parameter capturing the reciprocal of the minimum level of income compatible with positive welfare; specifically, $I(y,\bar{y}) = D(y,0)^{-1}$. We thank Gaurav Datt and Francois Bourguignon for pointing out the second and third equivalences, respectively.

(16) $$P(y,z) = \frac{1}{N}\sum_{i=1}^{N_z}\left(\frac{z}{y_i} - 1\right),$$

where $N_z \leq N$ corresponds to the number of individuals with incomes below z, and incomes are ordered such that $y_1 \leq y_2, \ldots, \leq y_N$. $P(y,z)$ takes on the value of zero when all incomes exceed the reference income level z. Like the welfare gap, $P(y,z)$ is based on the ratio z/y_i. However, unlike the welfare gap, it averages this ratio only across individuals below z. It therefore inherits the same attractive properties of subgroup decomposability and distribution sensitivity discussed above, although with the distribution-sensitivity properties satisfied only among individuals below z.[21] This poverty index is very closely related to that introduced in Bosmans, Esposito and Lambert (2011).[22]

The poverty index also inherits the same easily understood units of the welfare gap, although with a slight variant: $P(y,z)$ is the average *growth rate* needed to attain the reference income level, recognizing that zero growth is needed for people who are above the threshold.[23] Equivalently, we can write:

(17) $$P(y,z) = \left(\frac{N_z}{N}\right)\frac{1}{N_z}\sum_{i=1}^{N_z}\left(\frac{z}{y_i} - 1\right)$$

which is the fraction of the population below z times the average growth rate *among those below* z needed to achieve the reference income level. However, the poverty measure loses two attractive properties of the welfare gap. First, changing the reference income z no longer simply rescales the units of the poverty measure: it also changes the reference population over which the poverty measure is calculated and thus affects comparisons of income distributions. Second, like most poverty measures, $P(y,z)$ does not admit a simple analytical decomposition into the mean and a measure of inequality.

4. Empirical Illustration

4.1 Data

We illustrate the properties of the welfare gap and the mean ratio deviation using the global interpersonal income distribution from the World Bank's Poverty and Inequality Platform (PIP) (World Bank, 2022). PIP is a large compendium of over 2,000 household surveys covering 168

[21] We note that the poverty index $P(y,z)$ is similar to the poverty gap (a member of the Foster, Greer and Thorbecke (1984) class of poverty measures), but differs in one important respect: the poverty gap is based on the difference $z - y_i$, while $P(y,z)$ is based on the ratio z/y_i. Since the latter is convex while the former is not, $P(y,z)$ is distribution sensitive while the poverty gap is not. However, more convex members of the FGT class (with $\alpha > 1$) are subgroup decomposable and Pigou-Dalton sensitive.

[22] Specifically, Bosmans, Esposito and Lambert (2011) propose the index $\frac{1}{N}\sum_{i=1}^{N_z}\frac{z}{y_i}$. We subtract 1 in Equation (16) to ensure that the poverty index is continuous at the poverty line.

[23] We note the parallel between this interpretation and the Morduch (1998) interpretation of the Watts (1969) index as the average across individuals of the *continuously compounded* income growth rate required to exit poverty in T years.

countries and over 97 percent of the world's population.[24] For ease of reproducibility, the calculations reported here are performed on the "binned" version of the dataset created by Mahler ⓡ al. (2022), where the income distribution in each country-year is represented by data grouped into 1,000 bins of equal population size per country and year.[25]

Due to differences in available survey instruments, PIP combines surveys in which consumption is the main household-level measure of well-being (covering about three-quarters of the world's population) with surveys in which income is the measure of well-being. We follow the practice of the World Bank's global poverty estimates of ignoring the difference between income- and consumption-based surveys when combining distributions across countries. For terminological convenience, whenever possible we will refer to the measure of well-being from the survey as "income" irrespective of whether it is drawn from an income or a consumption survey. Income is measured in constant 2017 Purchasing Power Parity dollars ($PPP) per person per day. We bottom-code the data at $0.25 per day to prevent very low values of reported income from dominating our estimates of $W(y, z)$, by replacing all measured incomes $y_i < \$0.25$ with $y_i = \$0.25$ per day. We discuss the issue of bottom-coding in detail in Online Appendix B, showing that levels and trends of the welfare gap are robust within a plausible range of bottom-coding thresholds.

In the results below, we set the reference income level to $z = \$25$ per day. While the reference income level is just a scaling factor and does not impact the trends, rankings, or growth of the welfare measure, the value of $25 is salient for at least two reasons. First, $25 per day is close to the typical poverty line in high-income countries (Jolliffe ⓡ al., 2024). Second, a person living in a typical household in a middle-income country that is about to graduate into high-income status lives on an income of about $23 per day.[26]

4.2 Trends in the global welfare gap

Figure 1A reports the global welfare gap from 1990 to 2019. Since it is a "gap" measure, a decline in the welfare gap represents an improvement in welfare. In 1990, the welfare gap was 11.1, meaning that incomes on average needed to increase roughly 11-fold to bring everyone around the world to $25 per day. By 2019, the average gap had more than halved to 5.0. Since the welfare gap is homogenous of degree minus one in income, this is roughly equal to the gain in welfare that would have occurred if the income of every person in the world had doubled over this period.

Figure 1B unpacks the global welfare gap into three components. Combining Equations (14) and (15), the global welfare gap can be expressed as follows:

[24] We rely on the "lined-up" version of the PIP data which interpolates/extrapolates between/beyond survey years assuming distribution-neutral growth to create a continuous annual dataset covering 1990-2019. For details on the methodology used to line up surveys to a common reporting year, see the PIP methodological handbook (https://datanalytics.worldbank.org/PIP-Methodology). For countries without any household survey data, it is assumed that their income distribution is equal to the population-weighted average income distribution in their geographical region.

[25] The dataset is available at https://datacatalog.worldbank.org/search/dataset/0064304/1000-Binned-Global-Distribution (version September 2022).

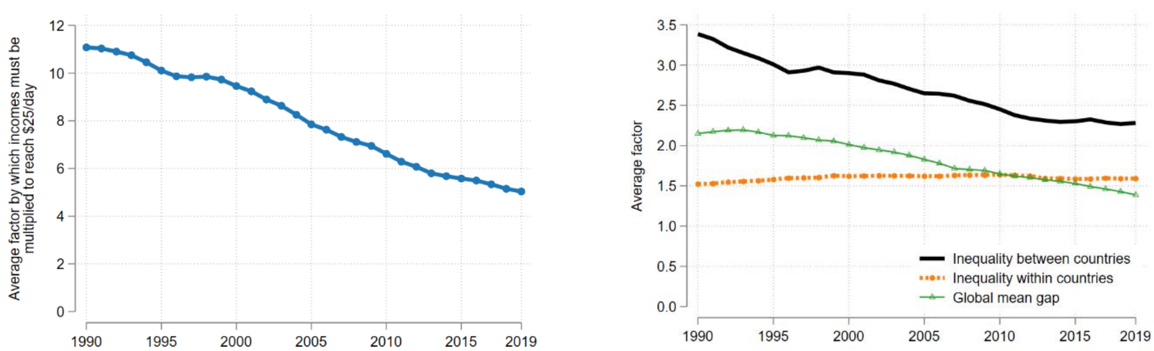
[26] See the working paper version of this paper (Kraay ⓡ al. 2023), Section 4.4 for details.

$$(18) \qquad \underbrace{W(y_t, z)}_{Welfare\ gap} = \underbrace{\frac{z}{\bar{y}_t}}_{global\ mean\ gap} \times \underbrace{I_{btw}(y_t, \bar{y}_t)}_{\substack{between-country \\ inequality}} \times \underbrace{I_{wth}(y_t, \bar{y}_t)}_{\substack{within-country \\ inequality}}$$

where I_{btw} and I_{wth} are defined as in Equation (15). In 1990, global mean income was roughly $\bar{y} = \$11.6$ per day, implying 2.1-fold shortfall from the $25 per day prosperity standard. Between 1990 and 2019, global mean income increased to $18 per day, resulting in a 1.4-fold shortfall from the same global prosperity standard in 2019 (green line in Figure 1B). Between-country inequality in 1990 was 3.4, indicating that the population-weighted average factor by which country mean income needed to be multiplied to reach global mean income was 3.4. By 2019, between-country inequality had declined to 2.3 (solid black line in Figure 1B). Finally, within-country inequality, measured as the weighted average factor by which individual incomes in each country need to be multiplied to reach their corresponding country average, remained largely unchanged during this period (orange dashed line in Figure 1B).

Figure 1: Global trends in $W(y, z)$

Panel A: Global welfare gap **Panel B: Components of the global welfare gap**



Source: Authors' calculation using Poverty and Inequality Platform (September 2022).

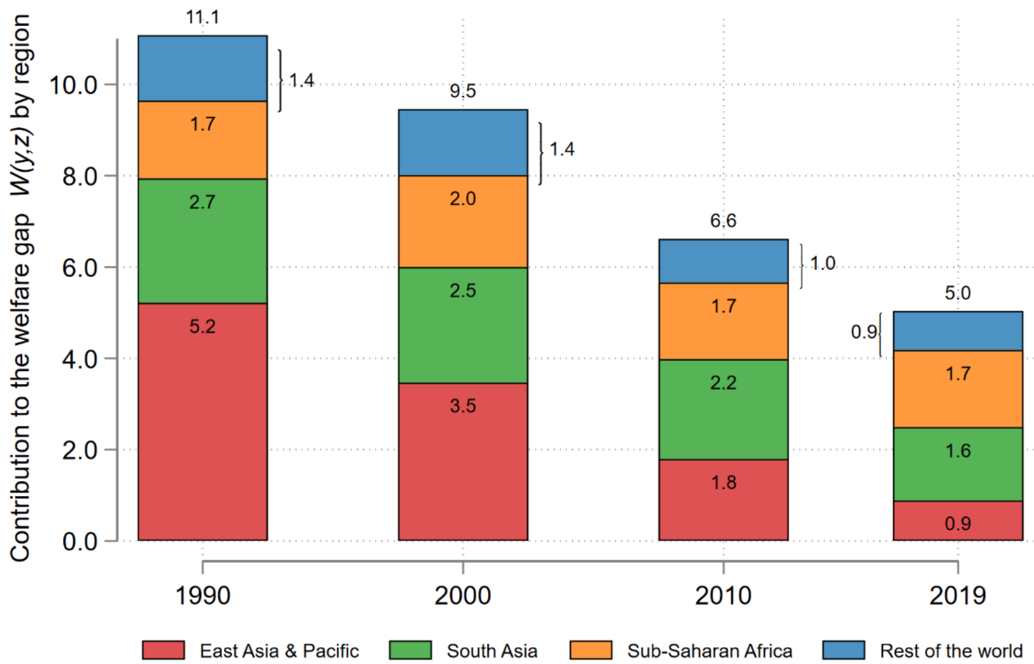
Notes: Panel A reports the global welfare gap $W(y, z)$ over the period 1990-2019 using z=$25 per person per day in 2017 $PPP as the reference income level. Panel B reports the three components of the global welfare gap, the product of which corresponds to the overall welfare gap shown in Panel A. Note that lower values of the welfare gap and its components imply higher welfare.

The global welfare gap in Figure 1A is the product of these three terms. As a result, the average annual rate of decline of the global welfare gap between 1990 and 2019 of 2.7 percent can be additively decomposed into (a) an average annual increase in global mean income of 1.5 percent; (b) an average annual reduction in between-country inequality of 1.4 percent; and (c) an average annual increase in within-country inequality of 0.2 percent.[27]

[27] The pattern of rising global average income and declining global interpersonal income inequality is not new. Mahler et al., (2022) and Deaton (2021) report similar patterns using the Gini index as a measure of inequality, while Lakner and Milanovic (2016) show that the decline in global inequality between 1988 and 2008 was primarily driven by the reductions in income gaps between countries. One key advantage of our measure is to allow for the joint analysis of

The decline (i.e., improvement) in the global welfare gap $W(y,z)$ of 2.7 percent per year is much larger than the improvement in global average income of 1.5 percent per year, reflecting an accompanying decline in global inequality $I(y,\bar{y})$. This pattern has largely been driven by rapid growth in East and South Asia, which in turn is due to the rapid economic growth of two large and initially poor countries, namely China and India. The population-weighted subgroup decomposability of the welfare gap allows us to document this in Figure 2.

Figure 2: Regional contributions to the global welfare gap



Source: Authors' calculation using Poverty and Inequality Platform (September 2022).

Notes: This figure reports the decomposition of the global welfare gap, $W(y,z)$, into regional contributions using the prosperity threshold of \$25 per person per day in 2017 \$PPP.

The largest contributors to the welfare gap are the 5 billion people in East Asia & Pacific, South Asia, and Sub-Saharan Africa, contributing more than 80% of the global welfare index in 1990 and in 2019. The (population-weighted) contribution of East Asia & Pacific to the global welfare gap declined from 5.2 to 0.9 during this period, while the corresponding figures were 2.7 and 1.6

trends in welfare, average income, and within- and between-group inequality, in a coherent and easily interpretable framework.

for South Asia.[28] In contrast, it is discouraging to note that the absolute contribution of Sub-Saharan Africa to the global welfare gap has not declined since 1990 and its share of the global welfare gap has actually increased from 15 percent in 1990 to 34 percent in 2019. This reflects the very high value of $W = 11.7$ in 2019 for Sub-Saharan Africa, implying that incomes in the region need to increase close to 12-fold on average to reach the reference income level of $25 per person per day. This is higher than the global welfare gap of 11.1 in 1990.

The distribution sensitivity of $W(y,z)$ can also substantially influence cross-country comparisons of welfare. Since the welfare gap is larger for more unequal countries, a country with higher average income may nevertheless have a higher welfare gap than another country with a lower average income but also lower inequality. For instance, although mean income in the United States in 2019 was 42 percent larger than in France, France's welfare gap was substantially lower than that of the United States (Table 2). This is because inequality in the United States, as measured by the mean ratio deviation, was three times as high as inequality in France. Holding fixed the distribution of income, mean income in the United States would have to be more than twice as high as it currently is (or three times that of France's) to achieve a welfare gap as low as France's. This corresponds to a sizeable "inequality penalty" for the United States relative to France.

Table 2: Comparing France and the United States using the welfare gap, mean income, and mean ratio deviation in 2019.

	France	United States
Welfare gap, $W(y,z)$	0.58	1.22
Mean income (per person/day)	58.9	83.6
Mean ratio deviation, $I(y,\bar{y})$	1.36	4.09

Notes: This table reports the estimates of $W(y,z)$, $I(y,\bar{y})$, and mean income for France and the United States. The welfare gap is reported using the $25 per person per day threshold in 2017 $PPP.

5. Conclusions

Widely used welfare and poverty measures such as mean income and the poverty headcount are not distribution sensitive. Despite widespread agreement that inequality matters for welfare comparisons, existing distribution sensitive welfare measures are much less frequently used in academic, policy, or public discourse. A likely reason for this is their relative complexity and lack of intuitive units.

To address this shortcoming, simpler distribution sensitive welfare measures are needed. Identifying and discriminating among such measures requires more than subjective assertions of simplicity. In this paper, we develop a novel axiomatic framework that proposes a set of well-defined simplicity properties and explores the restrictions they place on welfare measures. This approach exposes sharp trade-offs between simplicity properties satisfied by a welfare measure and the ethical properties of the welfare ordering it represents, as well as equally sharp trade-offs between different simplicity properties.

[28] The welfare gap $W(y,z)$ declined from 17.2 to 3.2 in East Asia & Pacific, and from 12.7 to 6.7 in South Asia.

Applying our approach leads to a new welfare measure, namely the welfare gap – the average factor by which incomes must be multiplied to reach a reference income level z. The welfare gap is distribution sensitive and is simple in the formal sense that it is uniquely characterized by a small set of simplicity properties. Given the limited uptake of existing distribution sensitive measures in policy discourse, we believe it is well worth pursuing alternative distribution sensitive measures that can readily be communicated to wider audiences and, therefore, may have greater policy impact. The recent adoption of the welfare gap by the World Bank as its headline measure of progress towards "shared prosperity," one of its two twin goals (World Bank, 2015), indicates some recognition of this view and is a meaningful step in this direction.

References

Alkire, S., Roche, J. M., Ballon, P., Foster, J., Santos, M. E., & Seth, S. (2015). *Multidimensional poverty measurement and analysis*. Oxford University Press, USA.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory*, *2*(3), 244-263.

Atkinson, A. B. (2016). *Monitoring global poverty. Report of the commission on global poverty. World Bank Group, Washington, DC*.

Bosmans, K., Esposito, L., & Lambert, P. (2011). New perspectives on a more-or-less familiar poverty index. *Social Choice and Welfare*, *36*(2), 241-247.

Chakravarty, S. R. (1983). A new index of poverty. Mathematical Social Sciences, 6(3), 307-313.

Chakravarty, S. R. (2009). Equity and efficiency as components of a social welfare function. *International Journal of Economic Theory*, *5*(2), 181-199.

Clark, S., Hemming, R., & Ulph, D. (1981). On indices for the measurement of poverty. *The Economic Journal*, *91*(362), 515-526.

Cowell, F., & Ebert, U. (2004). Complaints and inequality. *Social Choice and Welfare*, *23*(1), 71-89.

Cowell, F. A. and E. Flachaire (2021). Inequality measurement: Methods and data. In K. Zimmerman and E. Sierminska (Eds.), Handbook of Labor, Human Resources and Population Economics.

Cowell, F. A., and Victoria-Feser, M.-P. (1996). Robustness Properties of Inequality Measures. *Econometrica*, 64(1), 77–101.

Dalton, H. (1920). The Measurement of Inequality of Incomes. *The Economic Journal*, 30(119), 348-361.

Decerf, B. (2021). Combining absolute and relative poverty: income poverty measurement with two poverty lines. *Social Choice and Welfare*, *56*(2), 325-362.

Decerf, B. (2025). On the properties of the two main types of global poverty lines. *Journal of Development Economics*, *173*, 103396.

Deaton, A. (2021). COVID-19 and global income inequality (No. w28392). National Bureau of Economic Research.

Dollar, D., Kleineberg, T., & Kraay, A. (2015). Growth, Inequality and Social Welfare: Cross-Country Evidence. *Economic Policy.* 337-377.

Donaldson, D. and J.A. Weymark (1980). 'A single-parameter generalization of the Gini indices of inequality', Journal of Economic Theory, 22, 67–86.

Eden, M. (2022). Welfare Analysis with Heterogenous Risk Preferences. *Journal of Political Economy.* 128(12): 4574-4613.

Eden, M., and Freitas, L.M. (2024a). Income Anonymity. Working paper. https://drive.google.com/file/d/1lzG7-hYFTBR8nYMTk2H-3A55QM01hraI/view. *Accessed* 10-14-2024.

Eden, M., and Freitas, L.M. (2024b). Inequality Aversion in Developing Countries: Normative Guidelines Based on Income Anonymity. Working paper. https://drive.google.com/file/d/1mN9-sKemH6x5-DH3t6VnK7Y4hubgyFmB/view. *Accessed* 10-14-2024.

Fleurbaey, M. (2012). The importance of what people care about. *Politics, Philosophy & Economics*, *11*(4), 415-447.

Fleurbaey, M., & Maniquet, F. (2011). *A theory of fairness and social welfare* (Vol. 48). Cambridge University Press.

Fleurbaey, M., & Michel, P. (2001). Transfer principles and inequality aversion, with an application to optimal growth. *Mathematical Social Sciences*, *42*(1), 1-11.

Foster, J., Greer, J., & Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica: Journal of the Econometric Society*, 52(3): 761-766.

Foster, J. E., and M. Székely (2008). Is Economic Growth Good for the Poor? Tracking Low Incomes Using General Means. *International Economic Review.* 49(4): 1143-1172.

Hagenaars, A. (1987). A class of poverty indices. *International Economic Review*, 583-607.

Herforth, A., Venkat, A., Bai, Y., Costlow, L., Holleman, C., & Masters, W. A. (2022). Methods and options to monitor the cost and affordability of a healthy diet globally Background paper for The State of Food Security and Nutrition in the World 2022.

Hufe, P., Kanbur, R., & Peichl, A. (2022). Measuring unfair inequality: Reconciling equality of opportunity and freedom from poverty. *The Review of Economic Studies*, *89*(6), 3345-3380.

Jolliffeⓡ, D., Mahlerⓡ, D. G., Laknerⓡ, C., Atamanovⓡ, A., & Tetteh-Baah, S. K. (2024). Poverty and Prices: Assessing the Impact of the 2017 PPPs on the International Poverty Line and Global Poverty. *The World Bank Economic Review*, lhae035.

Kakwani, N. (1980). On a class of poverty measures. *Econometrica: Journal of the Econometric Society*, 437-446.

Kot, S. M., & Paradowski, P. R. (2022). The atlas of inequality aversion: theory and empirical evidence on 55 countries from the Luxembourg Income Study database. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *17*(2), 261-316.

Kraay, A. ⓡ Lakner, C. ⓡ Özler, B. ⓡ Decerf, B. ⓡ Jolliffe, D. ⓡ Sterck, O. ⓡ Yonzan, N. (2023). A New Distribution Sensitive Index for Measuring Welfare, Poverty, and Inequality. Policy Research Working Paper No. 10470, Washington, DC: World Bank.

Lakner, C., Milanovic, B.: Global income distribution: from the fall of the Berlin Wall to the great recession. World Bank Econ. Rev. 30(2), 203–232 (2016)

Mahler, D.G. ⓡ N. Yonzan ⓡ C. Lakner (2022). The Impact of COVID-19 on Global Inequality and Poverty. Policy Research Working Paper No. 10198, Washington, DC: World Bank.

Morduch, J. (1998). Poverty, economic growth, and average exit time. Economics Letters, 59(3), 385-390.

Neugschwender, J. (2020). Top and bottom coding at LIS (No. 9). LIS Technical Working Paper Series.

Piacquadio, P. G. (2017). A fairness justification of utilitarianism. *Econometrica*, *85*(4), 1261-1276.

Ravallion, M. (2016). Are the world's poorest being left behind? *Journal of Economic Growth*, 21, 139–164.

Ray, D., ⓡ Genicot, G. (2022). Measuring upward mobility (No. w29796). National Bureau of Economic Research.

Sen, A. (1976). Poverty: an ordinal approach to measurement. Econometrica: Journal of the Econometric Society, 219-231.

Shorrocks, Anthony F. "The class of additively decomposable inequality measures." *Econometrica: Journal of the Econometric Society* (1980): 613-625.

Sterck, O. (2024). Poverty without Poverty Line. *CSAE working paper.* WPS/2024-07.

Takayama, N. (1979). Poverty, income inequality, and their measures: Professor Sen's axiomatic approach reconsidered. Econometrica: Journal of the Econometric Society, 747-759.

Thon, D. (1979). On measuring poverty. Review of Income and Wealth, 25(4), 429-439.

Watts, H. W. (1969), An Economic Definition of Poverty, in D. P. Moynihan (ed.), On Understanding Poverty, New York: Basic Books.

World Bank (2015). A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data, and the Twin Goals. Policy Research Report. Washington, DC: World Bank.

World Bank (2022). Poverty and Inequality Platform (version 20220909_2017_01_02_PROD) [data set]. https://pip.worldbank.org/home.

Zheng, B. (1997). Aggregate poverty measures. Journal of economic surveys, 11(2), 123-16.

Appendix A: Summary of ethical axioms and simplicity properties satisfied by welfare measures discussed in this paper

	Axiom 1: Welfare Measure	Axiom 2: Pigou-Dalton Sensitive	SP1: Subgroup Decomposable	SP2: Strong Intuitive Units	SP3: Weak Intuitive Units	SP4: At Most One Income Parameter	SP5: Invariance to Income Parameters	SP6: Normalized Income Units	SP7: Increasing Welfare Measure	SP8: Non-Negative Values	SP9: Invariance to Income Units
Welfare Gap	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y
Mean	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N
Atkinson Index	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N
$M^1(y,z) = \frac{1}{N(y)}\sum_{i=1}^{N(y)}(y_i - \bar{z}) = \bar{y} - \bar{z}$	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N
$M^2(y,z) = \left(\frac{1}{N(y)}\sum_{i=1}^{N(y)}(y_i + \bar{z})^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}$	Y	Y	N	Y	Y	Y	N	N	Y	Y	N
$M^3(y,z) = Sen(y) = \bar{y} \times (1 - Gini(y))$	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N
$M^4(y) = \frac{1}{N(y)}\sum_{i=1}^{N(y)}(y_i)^{1/3}$	Y	Y	Y	N	N	Y	Y	N	Y	Y	N
$M^5(y,z) = (\bar{z})^{2/3} \times \frac{1}{N}\sum_{i=1}^{N}(y_i)^{1/3}$	Y	Y	Y	N	N	Y	Y	N	Y	Y	N

Online Appendix A: Proofs

We start with a preliminary Lemma.

Lemma 1: Take any parametric welfare measure M that satisfies SP1 and SP3 and any $z \in P$. For all $k > 0$ we have that $M_{kz}$is a welfare measure with the same direction as M_z (either they are both increasing, or both decreasing).

Proof of Lemma 1: Consider any $k > 0$ and $z \in P$. As M satisfies SP3 (Weak Intuitive Units), there are two cases.

Consider first the case for which M has income units. As $z \in P$, measure M_z represents a welfare ordering (which satisfies the Monotonicity Axiom). By Homogeneity of degree one in income units, we have for *all* $y, y' \in Y$ *that*

$$M(y,z) \geq M(y',z) \iff M(ky,kz) \geq M(ky',kz)$$

This implies that measure M_{kz} is a welfare measure (M_{kz} satisfies the Monotonicity Axiom) and has the same direction as M_z.

Consider then the case for which M is the ratio of two incomes. By Definition 3, there exists some partition of its income parameters $z = (z^{ref}, z^*)$ and a parametric measure M^* that has income units such that we have either $M(y,z) = \frac{M^*(y,z^*)}{M^*(z^{ref},z^*)}$, or $M(y,z) = \frac{M^*(z^{ref},z^*)}{M^*(y,z^*)}$.

First, we consider $M^*(z^{ref}, z^*) = 0$. For the case $M(y,z) = \frac{M^*(y,z^*)}{M^*(z^{ref},z^*)}$, we cannot have $M^*(z^{ref}, z^*) = 0$ (otherwise $M(y,z) = \emptyset$ and thus M is not a welfare measure). For the case $M(y,z) = \frac{M^*(z^{ref},z^*)}{M^*(y,z^*)}$, if we have $M^*(z^{ref}, z^*) = 0$, then we have $M^*(kz^{ref}, kz^*) = 0$ by homogeneity of degree one in income units. Thus $M(y,z) = M(y,kz) = M(ky,kz) = 0$ for all $y \in Y$. This shows that measure M_{kz} represents the same welfare ordering as M_z, which satisfies the Monotonicity Axiom. We thus have that M_{kz} is a welfare measure with the same direction as M_z.

Next, we consider $M^*(z^{ref}, z^*) \neq 0$. By homogeneity of degree one in income units, we have $M^*(kz^{ref}, kz^*) = kM^*(z^{ref}, z^*)$. This implies that $M^*(kz^{ref}, kz^*) \neq 0$ because $M^*(z^{ref}, z^*) \neq 0$. Also, $M^*(kz^{ref}, kz^*)$ has the same sign as $M^*(z^{ref}, z^*)$. Thus, M_{kz} is a welfare measure with the same direction as M_z if ${M_{kz}}^*$ is a welfare measure with the same direction as ${M_z}^*$. This condition holds because M^* has income units (following the argument developed for the case where M has income units). This concludes the proof of Lemma 1.

Proof of Theorem 1

$\Leftarrow$ Sufficiency. The proof that any measure M that has the mathematical expression provided by the condition is a welfare measure and satisfies SP1 and SP2 is immediate and thus omitted.

⇒ Necessity. We prove that any welfare measure M that satisfies SP1 and SP2 must have the mathematical expression provided by the condition.

By recursive application of SP1, measure M is such that $M(y,z) = \frac{1}{N(y)} \sum_{i=1}^{N(y)} f(y_i, z)$ for all y ∈ Y and all z ∈ P. There remains to characterize the contribution function f. Take any z ∈ P. By SP2, M satisfies weak Normalization. For income distributions with a unique income, this implies that for all $a, b > 0$ *with* $b > a$ we have $|f(b, z) - f(a, z)| = b - a$.

There are two cases, either M is an increasing welfare measure or M is a decreasing welfare measure. Consider the case for which M is an increasing measure (which covers the case $M = M^+$ in the statement of Theorem 1). The monotonicity axiom implies that $f(b, z) - f(a, z) = b - a$ for all $a, b > 0$ *with* $b > a$. This equality in turns implies that $f(y_i, z) = y_i - B(z)$ where $B(z) \in R$. Finally, by SP2, M satisfies homogeneity of degree one in income units, which implies that for *all* $k > 0$, *all* $y_i > 0$ *and all* $z \in P$ *we have* $f(ky_i, kz) = kf(y_i, z)$ (By Lemma 1, $z \in P$ implies that M_{kz} is a welfare measure with the same direction as M_z). This implies that $B: P \rightarrow R$ is homogeneous of degree one in its argument z. The case for which M is a decreasing welfare measure (which covers the case $M = -M^+$ in the statement of Theorem 1) yields $f(y_i, z) = B(z) - y_i$ using the same reasoning. This concludes the proof of Theorem 1.

Proof of Theorem 2

⇐ Sufficiency. The proof that any measure M that has the mathematical expression provided by the condition is a welfare measure and satisfies SP1 and SP3 is immediate and thus omitted.

⇒ Necessity. We prove that any welfare measure M that satisfies SP1 and SP3 must have the mathematical expression provided by the condition.

Case 1: M has income units.

This case is identical to that in Theorem 1 above, as M satisfies SP1 and SP2. The proof, yielding Eq (1) in ***Theorem 2***, is thus omitted.

The non-trivial cases are those for which M is the ratio of two incomes: ***Cases 2*** & ***3*** below refer, respectively, to the first & second parts of ***Definition 3***, yielding Equations (11) & (12) in ***Theorem 2***.

If M is the ratio of two incomes then, by Definition 3, there exists some partition of its income parameters $z = (z^{ref}, z^*)$ and a parametric measure M^* that has income units such that we have either $M(y,z) = \frac{M^*(y,z^*)}{M^*(z^{ref},z^*)}$, or $M(y,z) = \frac{M^*(z^{ref},z^*)}{M^*(y,z^*)}$ (*for all* $y \in Y$ *and all* $z \in P$).

Case 2: $\boldsymbol{M(y,z) = \frac{M^*(y,z^*)}{M^*(z^{ref},z^*)}}$

As M^* satisfies homogeneity of degree one in income units, we have for *all* $k > 0$, *all* $y \in Y$ *and all* $z \in P$ *that* $M^*(ky, kz^*) = kM^*(y, z^*)$ and $M^*(kz^{ref}, kz^*) = kM^*(z^{ref}, z^*)$. Case 2 in turn implies, for *all* $k > 0$, *all* $y_i > 0$ *and all* $z \in P$, that $f(ky_i, kz) = f(y_i, z)$, i.e., function f is homogeneous of degree zero (By Lemma 1, $z \in P$ implies that M_{kz} is a welfare measure with the same direction as M_z).

SP1 (subgroup decomposability) implies that

$$\frac{1}{N(y)}\sum_{i=1}^{N(y)} f(y_i, z) = \frac{M^*(y, z^*)}{M^*(z^{ref}, z^*)} \qquad for\ all\ y \in Y\ and\ all\ z \in P$$

where $f: R \times P \to R$.

We can rewrite the last equation as follows:

$$M^*(y, z^*) = M^*(z^{ref}, z^*) \times \frac{1}{N(y)}\sum_{i=1}^{N(y)} f(y_i, z) \qquad for\ all\ y \in Y\ and\ all\ z \in P \qquad \textbf{(Eq A.1)}$$

Take any $z \in P$. Consider two income distributions with N individuals each, $a\mathbb{1}_N$ and $b\mathbb{1}_N$, where $a, b > 0\ with\ b > a$. As M^* has income units, it satisfies Weak Normalization, therefore:

$|M^*(b\mathbb{1}_N, z^*) - M^*(a\mathbb{1}_N, z^*)| = b - a.$

Substituting for $M^*(y, z^*)$ using Eq (A.1) and taking means, we get

$\left|M^*(z^{ref}, z^*) \times \left(f(b, z) - f(a, z)\right)\right| = b - a$

which is valid for all $a, b > 0\ with\ b > a.$

The sign of the product inside the absolute value brackets is determined by whether (a) $M^*(z^{ref}, z^*)$ is positive or negative, and (b) M is an increasing or decreasing measure.[29] If M is increasing (decreasing), by monotonicity, f is increasing (decreasing) in income y_i, implying that the second term is positive (negative). The following table summarizes the sign of $\left(M^*(b\mathbb{1}_N, z^*) - M^*(a\mathbb{1}_N, z^*)\right)$:

Table A.1: Sign of $\left(M^*(b\mathbb{1}_N, z^*) - M^*(a\mathbb{1}_N, z^*)\right)$ and implied mathematical expression for measure M.

IF	M is an increasing measure	M is a decreasing measure
$M^*(z^{ref}, z^*) > 0$	$\boldsymbol{b - a > 0}$ $\Rightarrow M(y,z) = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)}$	$\boldsymbol{a - b < 0}$ $\Rightarrow M(y,z) = \frac{B(z^*) - \bar{y}}{B(z^*) - \bar{z}^{ref}}$
$M^*(z^{ref}, z^*) < 0$	$\boldsymbol{a - b < 0}$ $\Rightarrow M(y,z) = \frac{B(z^*) - \bar{y}}{B(z^*) - \bar{z}^{ref}}$	$\boldsymbol{b - a > 0}$ $\Rightarrow M(y,z) = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)}$

[29] By definition, a welfare measure is either increasing for all its parameters values, or decreasing for all its parameters values.

As we can see from the table, there are four sub-cases of ***Case 2***. (If $M^*(z^{ref}, z^*) = 0$, then M is not well-defined and thus cannot represent a welfare ordering ($z \notin P$).) It will suffice to prove only one subcase in detail, as the other three cases follow using the same reasoning.

Sub-case 2.1: M is an increasing measure and $M^*(z^{ref}, z^*) > 0$

From Table A.1 above, $M^*(z^{ref}, z^*) \times \left(f(b, z) - f(a, z)\right) = b - a$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^1$, where

$$f^1(y_i, z) = \frac{y_i - B(z^*)}{M^*(z^{ref}, z^*)} \quad for\ all\ y_i > 0$$

where $B(z^*) \in R$.[30] As function M^* is homogeneous of degree 1 and function f^1 is homogeneous of degree zero, function B must be homogeneous of degree one in z^*. Substituting for f^1 in Eq (A.1), we get that $M^*(y, z^*) = \frac{1}{N(y)} \sum_{i=1}^{N} y_i - B(z^*) = \bar{y} - B(z^*)$ for all $y \in Y$. Therefore, we have $M^*(z^{ref}, z^*) = \bar{z}^{ref} - B(z^*)$. Since, in this sub-case, $M^*(z^{ref}, z^*) > 0$, we have $\bar{z}^{ref} > B(z^*)$.

Thus, we have shown that when M is an increasing measure and $M^*(z^{ref}, z^*) > 0$, $M(y, z) = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)}$, $B: P \to R$ is *homogeneous of degree one,* and $\bar{z}^{ref} > B(z^*)$ for all $z \in P$. To complete the part of the proof for Case 2 in Theorem 2 for which M is an increasing measure, we still need to show that $\bar{z}^{ref} > B(z^*)$ when $M^*(z^{ref}, z^*) < 0$. We demonstrate this in sub-case 2.2 below.

Sub-case 2.2: M is an increasing measure and $M^*(z^{ref}, z^*) < 0$

From Table A.1 above, $M^*(z^{ref}, z^*) \times \left(f(b, z) - f(a, z)\right) = a - b$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^2$, where

$$f^2(y_i, z) = \frac{B(z^*) - y_i}{M^*(z^{ref}, z^*)} \quad for\ all\ y_i > 0$$

where $B(z^*) \in R$ and B is homogeneous of degree one in z^*. Using the same reasoning in Sub-case 2.1, $M^*(z^{ref}, z^*) = B(z^*) - \bar{z}^{ref}$, which implies that $\bar{z}^{ref} > B(z^*)$.

Thus, we have shown that when M is an increasing measure and $M^*(z^{ref}, z^*) < 0$, $M(y, z) = \frac{B(z^*) - \bar{y}}{B(z^*) - \bar{z}^{ref}}$, $B: P \to R$ is *homogeneous of degree one,* and $\bar{z}^{ref} > B(z^*)$ for all $z \in P$.

As $\frac{B(z^*) - \bar{y}}{B(z^*) - \bar{z}^{ref}} = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)}$, ***sub-cases 2.1 & 2.2 complete the proof for Case 2 in Theorem 2, for which*** $\bar{z}^{ref} > B(z^*)$ ***for all z ∈ P (M is an increasing measure).***

The following two sub-cases cover Case 2 in Theorem 2 when M is a decreasing measure, for which $\bar{z}^{ref} < B(z^*)$ for all $z \in P$.

---

[30] As function M^* does not admit z^{ref} among its income parameters, we must have $B(z^*)$, not $B(z^{ref}, z^*)$.

Sub-case 2.3: M is a decreasing measure and $M^*(z^{ref}, z^*) > 0$

From Table A.1 above, $M^*(z^{ref}, z^*) \times \big(f(b, z) - f(a, z)\big) = a - b$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^2$, where

$$f^2(y_i, z) = \frac{B(z^*) - y_i}{M^*(z^{ref}, z^*)} \quad for\ all\ y_i > 0$$

where $B(z^*) \in R$ and B is homogeneous of degree one in z^*. Using the same reasoning as above, $M^*(z^{ref}, z^*) = B(z^*) - \bar{z}^{ref}$, which implies that $\bar{z}^{ref} < B(z^*)$.

Thus, we have shown that when M is a decreasing measure and $M^*(z^{ref}, z^*) > 0$, $M(y, z) = \frac{B(z^*) - \bar{y}}{B(z^*) - \bar{z}^{ref}} = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)}$, $B: P \to R$ is *homogeneous of degree one,* and $\bar{z}^{ref} > B(z^*)$ for all $z \in P$. Finally, to complete the part of the proof for Case 2 in Theorem 2 for which M is a decreasing measure, we need to show that $\bar{z}^{ref} < B(z^*)$ when $M^*(z^{ref}, z^*) < 0$. We demonstrate this in sub-case 2.4 below.

Sub-case 2.4: M is a decreasing measure and $M^*(z^{ref}, z^*) < 0$

From Table A.1 above, $M^*(z^{ref}, z^*) \times \big(f(b, z) - f(a, z)\big) = b - a$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^1$, where

$$f^1(y_i, z) = \frac{y_i - B(z^*)}{M^*(z^{ref}, z^*)} \quad for\ all\ y_i > 0$$

where $B(z^*) \in R$ and B is homogeneous of degree one in z^*. Using the same reasoning as above, $M^*(z^{ref}, z^*) = \bar{z}^{ref} - B(z^*)$, which implies that $\bar{z}^{ref} < B(z^*)$.

Thus, we have shown that when M is a decreasing measure and $M^*(z^{ref}, z^*) < 0$, $M(y, z) = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)}$, $B: P \to R$ is *homogeneous of degree one,* and $\bar{z}^{ref} < B(z^*)$ for all $z \in P$.

Sub-cases 2.3 & 2.4 cover ***Case 2*** in ***Theorem 2***, for which $\bar{z}^{ref} < B(z^*)$ for all $z \in P$ ***(M is a decreasing measure)***.

Sub-cases 2.1-2.4 complete the proof of Case 2 in Theorem 2.

Case 3: $M(y, z) = \frac{M^*(z^{ref}, z^*)}{M^*(y, z^*)}$

As M^* satisfies homogeneity of degree one in income units, we have for *all* $k > 0$*, all* $y_i > 0$ *and all* $z \in P$ *that* $M^*(ky, kz^*) = kM^*(y, z^*)$ and $M^*(kz^{ref}, kz^*) = kM^*(z^{ref}, z^*)$. This in turns implies for *all* $k > 0$*, all* $y_i > 0$ *and all* $z \in P$ that $f(ky_i, kz) = f(y_i, z)$, i.e., function f is homogeneous of degree zero (By Lemma 1, $z \in P$ implies that M_{kz} is a welfare measure with the same direction as M_z).

SP1 (subgroup decomposability) implies that

$$\frac{1}{N(y)}\sum_{i=1}^{N(y)} f(y_i,z) = \frac{M^*(z^{ref},z^*)}{M^*(y,z^*)} \qquad for\ all\ y \in Y\ and\ all\ z \in P$$

where $f: R \times P \rightarrow R$. We can rewrite last equation as

$$M^*(y,z^*) = \frac{M^*(z^{ref},z^*)}{\frac{1}{N(y)}\sum_{i=1}^{N(y)} f(y_i,z)} \qquad for\ all\ y \in Y\ and\ all\ z \in P \qquad \textbf{(Eq A.2)}$$

Take any $z \in P$. Consider two income distributions with N individuals each, $a\mathbb{1}_N$ and $b\mathbb{1}_N$, where $a, b > 0\ with\ b > a$. As M^* has income units, it satisfies Weak Normalization, therefore:

$|M^*(b\mathbb{1}_N, z^*) - M^*(a\mathbb{1}_N, z^*)| = b - a.$

Substituting for $M^*(y,z^*)$ using Eq (A.2) and taking means, we get

$\left|\frac{M^*(z^{ref},z^*)}{f(b,z)} - \frac{M^*(z^{ref},z^*)}{f(a,z)}\right| = b - a$

which is valid for all $a, b > 0\ with\ b > a$. Clearly, last equation cannot hold if $M^*(z^{ref}, z^*) = 0$, so we must have $M^*(z^{ref}, z^*) \neq 0$.

The sign of the mathematical expression inside the absolute value brackets is determined by whether (a) $M^*(z^{ref}, z^*)$ is positive or negative, and (b) M is an increasing or decreasing measure. [31] If M is increasing (decreasing), by monotonicity, f is increasing (decreasing) in income y_i, implying that the second term is positive (negative). The following table summarizes the sign of $\left(M^*(b\mathbb{1}_N, z^*) - M^*(a\mathbb{1}_N, z^*)\right)$:

Table A.2: Sign of $\left(M^*(b\mathbb{1}_N, z^*) - M^*(a\mathbb{1}_N, z^*)\right)$ and implied mathematical expression for measure M.

IF	M is an increasing measure	M is a decreasing measure
$M^*(z^{ref}, z^*) > 0$	$\boldsymbol{a - b < 0}$ $\Rightarrow Impossible$	$\boldsymbol{b - a > 0}$ $\Rightarrow M(y,z) = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)}$
$M^*(z^{ref}, z^*) < 0$	$\boldsymbol{b - a > 0}$ $\Rightarrow Impossible$	$\boldsymbol{a - b < 0}$ $\Rightarrow M(y,z) = \frac{Harm\left(B(z^*)\mathbb{1}_{N(z^{ref})} - z^{ref}\right)}{Harm\left(B(z^*)\mathbb{1}_{N(y)} - y\right)}$

[31] By definition, a welfare measure is either increasing for all its parameters values, or decreasing for all its parameters values.

Analogous to ***Case 2***, there are four sub-cases of ***Case 3***. (We cannot have $M^*(z^{ref}, z^*) = 0$, as shown above) Similarly, it will suffice to prove only one subcase in detail, as the other three cases follow using the same reasoning.

Sub-case 3.1: M is an increasing measure and $M^*(z^{ref}, z^*) > 0$

From Table A.2 above, $\frac{M^*(z^{ref},z^*)}{f(b,z)} - \frac{M^*(z^{ref},z^*)}{f(a,z)} = a - b < 0$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^3$, where

$$f^3(y_i, z) = \frac{M^*(z^{ref}, z^*)}{B(z^*) - y_i} \quad for\ all\ y_i > 0$$

where $B(z^*) \in R$.[32] When $B(z^*) > 0$, function f^3 is undefined for $y_i = B(z^*)$ and thus measure M does not represent a *complete* ordering (Definition 1). This implies that $B: P \to R_{\leq 0}$.

As function M^* is homogeneous of degree 1 and function f^3 is homogeneous of degree zero, function B must be homogeneous of degree one in z^*. Entering the mathematical expression for the contribution function f^3 in Eq (A.2), we get that

$$M^*(y, z^*) = \frac{1}{\frac{1}{N(y)}\sum_i \frac{1}{B(z^*) - y_i}} = Harm\big(B(z^*)\mathbb{1}_{N(y)} - y\big)$$

for all $y \in Y$. Therefore, we have $M^*(z^{ref}, z^*) = Harm\left(B(z^*)\mathbb{1}_{N(z^{ref})} - z^{ref}\right)$. As $M^*(z^{ref}, z^*) > 0$ in this sub-case, $Harm\left(B(z^*)\mathbb{1}_{N(z^{ref})} - z^{ref}\right) > 0$, which is impossible given that $z_j^{ref} > 0$ for all j and $B(z^*) \leq 0$.

Thus, we have shown that when M is an increasing measure and $M^*(z^{ref}, z^*) > 0$, it is impossible for $M(y, z) = \frac{Harm\left(B(z^*)\mathbb{1}_{N(z^{ref})} - z^{ref}\right)}{Harm\left(B(z^*)\mathbb{1}_{N(y)} - y\right)} = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)}$, where function $B: P \to R_{\leq 0}$ is *homogeneous of degree one*. To complete the part of the proof for Case 3 in Theorem 2 for which M is an increasing measure, we still need to show the same impossibility when $M^*(z^{ref}, z^*) < 0$. We demonstrate this in sub-case 3.2 below.

Sub-case 3.2: M is an increasing measure and $M^*(z^{ref}, z^*) < 0$

From Table A.2 above, $\frac{M^*(z^{ref},z^*)}{f(b,z)} - \frac{M^*(z^{ref},z^*)}{f(a,z)} = b - a > 0$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^4$, where

$$f^4(y_i, z) = \frac{M^*(z^{ref}, z^*)}{y_i - B(z^*)} \quad for\ all\ y_i > 0$$

[32] As function M^* does not admits z^{ref} among its income parameters, we must have $B(z^*)$, not $B(z^{ref}, z^*)$.

where $B(z^*) \in R$.[33] When $B(z^*) > 0$, function f^4 is undefined for $y_i = B(z^*)$ and thus measure M does not represent a *complete* ordering (Definition 1). This implies that $B: P \to R_{\leq 0}$.

As function M^* is homogeneous of degree 1 and function f^4 is homogeneous of degree zero, function B must be homogeneous of degree one in z^*. Using the same reasoning in sub-case 3.1, we have $M^*(z^{ref}, z^*) = Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)$. As $M^*(z^{ref}, z^*) < 0$ in this sub-case, $Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right) < 0$, which is impossible given that $z_j^{ref} > 0$ for all j and $B(z^*) \leq 0$.

Thus, we have shown that when M is an increasing measure and $M^*(z^{ref}, z^*) < 0$, it is impossible for $M(y,z) = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)}$, where function $B: P \to R_{\leq 0}$ is *homogeneous of degree one*.

This completes the part of the proof for Case 3 in Theorem 2 for which M is an increasing measure, and shows that it is impossible for $M(y,z) = \frac{M^*(z^{ref}, z^*)}{M^*(y, z^*)}$ if M is an increasing measure.

Sub-case 3.3: M is a decreasing measure and $M^*(z^{ref}, z^*) > 0$

From Table A.2 above, $\frac{M^*(z^{ref}, z^*)}{f(b,z)} - \frac{M^*(z^{ref}, z^*)}{f(a,z)} = b - a > 0$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^4$, where

$$f^4(y_i, z) = \frac{M^*(z^{ref}, z^*)}{y_i - B(z^*)} \quad for\ all\ y_i > 0$$

where $B(z^*) \in R$.[34] When $B(z^*) > 0$, function f^4 is undefined for $y_i = B(z^*)$. This implies that $B: P \to R_{\leq 0}$.

Using the same reasoning in sub-case 3.2, we have $M^*(z^{ref}, z^*) = Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)$. As $M^*(z^{ref}, z^*) > 0$ in this sub-case, $Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right) > 0$, given that $z_j^{ref} > 0$ for all j and $B(z^*) \leq 0$.

Thus, we have shown that when M is a decreasing measure and $M^*(z^{ref}, z^*) > 0$, $M(y,z) = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)}$, where function $B: P \to R_{\leq 0}$ is *homogeneous of degree one*. To complete the proof of Theorem 2, we finally need to show that $M(y,z) = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)}$ when M is a decreasing measure and $M^*(z^{ref}, z^*) > 0$. We demonstrate this in sub-case 3.4 below.

[33] As function M^* does not admits z^{ref} among its income parameters, we must have $B(z^*)$, not $B(z^{ref}, z^*)$.
[34] As function M^* does not admits z^{ref} among its income parameters, we must have $B(z^*)$, not $B(z^{ref}, z^*)$.

Sub-case 3.4: M is a decreasing measure and $M^*(z^{ref}, z^*) < 0$

From Table A.2 above, $\frac{M^*(z^{ref},z^*)}{f(b,z)} - \frac{M^*(z^{ref},z^*)}{f(a,z)} = a - b < 0$, which holds for all $a, b > 0$ *with* $b > a$. Therefore, it must be the case that $f = f^3$, where

$$f^3(y_i, z) = \frac{M^*(z^{ref}, z^*)}{B(z^*) - y_i} \quad for\ all\ y_i > 0$$

where $B(z^*) \in R$.[35] When $B(z^*) > 0$, function f^3 is undefined for $y_i = B(z^*)$ and thus measure M does not represent a *complete* ordering (Definition 1). This implies that $B: P \to R_{\leq 0}$.

Using the same reasoning in sub-case 3.1, we have $M^*(z^{ref}, z^*) = Harm\left(B(z^*)\mathbb{1}_{N(z^{ref})} - z^{ref}\right)$. As $M^*(z^{ref}, z^*) < 0$ in this sub-case, $Harm\left(B(z^*)\mathbb{1}_{N(z^{ref})} - z^{ref}\right) < 0$, given that $z_j^{ref} > 0$ for all j and $B(z^*) \leq 0$.

Thus, we have shown that when M is a decreasing measure and $M^*(z^{ref}, z^*) < 0$, $M(y, z) = \frac{Harm\left(B(z^*)\mathbb{1}_{N(z^{ref})} - z^{ref}\right)}{Harm\left(B(z^*)\mathbb{1}_{N(y)} - y\right)} = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)}$, where function $B: P \to R_{\leq 0}$ is *homogeneous of degree one*.

This completes the part of the proof for Case 3 in Theorem 2 for which M is a decreasing measure, and shows that if $M(y, z) = \frac{M^*(z^{ref},z^*)}{M^*(y,z^*)}$, then M has to be a decreasing measure.

This completes the proof of Theorem 2.

Proof of Theorem 3

⇐ Sufficiency. Proof omitted.

⇒ Necessity. We prove that $M(y, z_1) = \frac{\bar{y}}{z_1}$ *and* $M(y, z_1) = \frac{z_1}{Harm(y)}$ are the only welfare measures that satisfy SP1, SP3 and SP4.

By SP4, we have that $z = z_1 > 0$, namely that z admits at most one income parameter. By Theorem 2, we have that $M(y, z_1)$ takes one of three mathematical expressions, which we consider in turn.

Case 1: $M = M^+$ *or* $M = -M^+$ for some increasing measure M^+ such that for all $y \in Y$ and all $z \in P$

$$M^+(y, z) = \bar{y} - B(z) \qquad (10)$$

where $B: P \to R$ that is *homogeneous of degree one*.

This directly yields Case 1 in Theorem 3 when switching z for z_1.

[35] As function M^* does not admits z^{ref} among its income parameters, we must have $B(z^*)$, not $B(z^{ref}, z^*)$.

Case 2: for some partition of income parameters $z = (z^{ref}, z^*)$ we have for all y ∈ Y and all $z \in P$ that

$$M(y,z) = \frac{\bar{y} - B(z^*)}{\bar{z}^{ref} - B(z^*)} = \frac{Mean\big(y - B(z^*)\mathbb{1}_{N(y)}\big)}{Mean\big(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\big)} \qquad (11)$$

where function $B: P \to R$ is *homogeneous of degree one* and either $\bar{z}^{ref} > B(z^*)$ for all $z \in P$ or $\bar{z}^{ref} < B(z^*)$ for all $z \in P$;

There are two possibilities for the partition: either $z^* = z_1$ and $z^{ref} = \emptyset$ or $z^{ref} = z_1$ and $z^* = \emptyset$. Consider first the case $z^* = z_1$ and $z^{ref} = \emptyset$. This case is such that $\bar{z}^{ref} = \emptyset$ and thus $Mean\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right) = \emptyset$. As a result, M is not a welfare measure, a contradiction. There remains the case $z^{ref} = z_1$ and $z^* = \emptyset$. Function B has no argument. However, function B must be homogeneous of degree one so that M is homogeneous of degree zero. The only possibility is $B(z^*) = 0$, which is trivially homogeneous of degree one. Finally, we get $M(y, z_1) = \frac{\bar{y}}{z_1}$ when observing that $Mean(z_1) = z_1$, which is such that $z_1 > B(z^*)$. This yields Case 2 in Theorem 3.

Case 3: for some partition of income parameters $z = (z^{ref}, z^*)$ we have for all y ∈ Y and all $z \in P$ that

$$M(y,z) = \frac{\frac{1}{N(y)}\sum_i \frac{1}{y_i - B(z^*)}}{\frac{1}{N(z^{ref})}\sum_j \frac{1}{z_j^{ref} - B(z^*)}} = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\big(y - B(z^*)\mathbb{1}_{N(y)}\big)} \qquad (12)$$

where function $B: P \to R_{\leq 0}$ is *homogeneous of degree one*.

The proof that Case 3 leads to $M(y, z_1) = \frac{z_1}{Harm(y)}$ follows the same reasoning as for Case 2 (observing that $Harm(z_1) = z_1$) and is thus omitted.

This concludes the proof of Theorem 3.

Proof of Theorem 4

The proof that the welfare gap (Statement 3) satisfies the definition for measure M given in statements 1 and 2 is omitted.

We prove that any measure M defined in statements 1 and 2 must satisfy Statement 3 (i.e., be the welfare gap).

As mean income is not Pigou-Dalton sensitive, we have by Theorem 2 that any welfare measure that satisfies SP1 and SP3 and is Pigou-Dalton sensitive is such that for some partition of income parameters $z = (z^{ref}, z^*)$ we have for all y ∈ Y and all $z \in P$ that

$$M(y,z) = \frac{\frac{1}{N(y)}\sum_i \frac{1}{y_i - B(z^*)}}{\frac{1}{N(z^{ref})}\sum_j \frac{1}{z_j^{ref} - B(z^*)}} = \frac{Harm\left(z^{ref} - B(z^*)\mathbb{1}_{N(z^{ref})}\right)}{Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)} \quad (12)$$

where function $B: P \to R_{\leq 0}$ is *homogeneous of degree one*.

Statement 1: Consider Eq (12). As $Harm(z^{ref}) > 0$ for all $(z^{ref}, z^*) \in P$, it is sufficient to prove that $B(z^*) = 0$ for all $(z^{ref}, z^*) \in P$. Assume to the contrary that $B(z^*) = -a$ for some $a > 0$ and some $(z^{ref}, z^*) \in P$. For any $k > 1$, we have $(kz^{ref}, kz^*) \in P$ and $B(kz^*) = -ka$ (since function B is homogeneous of degree one). However, for all $a > 0$ and $k > 1$ there exist two distributions $y, y' \in Y$ such that $Harm\left(y' + a\mathbb{1}_{N(y)}\right) > Harm\left(y + a\mathbb{1}_{N(y)}\right)$ but $Harm\left(y' + ka\mathbb{1}_{N(y)}\right) < Harm\left(y + ka\mathbb{1}_{N(y)}\right)$.[36] We get from Eq (12) that the ordering made by measure M between distributions y and y' changes between (z^{ref}, z^*) and (kz^{ref}, kz^*), which violates SP5. This shows that $B(z^*) = 0$ for all $(z^{ref}, z^*) \in P$. We obtain from Eq (12) that $M(y, z_1) = \frac{z_1}{Harm(y)}$ where $z_1 = Harm(z^{ref})$, the desired result.

Statement 2: The measure M defined in Eq (12) relies on a measure M^* that has income units defined as $M^*(y, z^*) = Harm\left(y - B(z^*)\mathbb{1}_{N(y)}\right)$. Measure M^* has normalized income units only if $B(z^*) = 0$ for all $(z^{ref}, z^*) \in P$.[37] As in the proof of Statement 3, we obtain from Eq (12) that $M(y, z_1) = \frac{z_1}{Harm(y)}$ where $z_1 = Harm(z^{ref})$, the desired result.

This concludes the proof of Theorem 4.

[36] For instance, we obtain the desired inequalities when defining distributions $y = (1, x)$ and $y' = (2,2)$ and selecting $x \in \left(\frac{3ak+2}{ak}, \frac{3a+2}{a}\right)$.

[37] Indeed, if B(z*)≠0, then we have for any $a > 0$ that $M^*(a\mathbb{1}_n, z^*) = Harm(a\mathbb{1}_n - B(z^*)\mathbb{1}_n) = a - B(z^*) \neq a$, which violates Strong Normalization.

Online Appendix B: Sensitivity to very low incomes

The welfare gap and its related inequality and poverty indices are based on the average of z/y_i, meaning that they are undefined when the data include values of $y_i = 0$. They are also highly sensitive to low values of y_i, becoming arbitrarily large as the lowest value of y_i in the survey approaches zero from above. The same is true for some other measures such as the Watts index, the Atkinson Index with $\varepsilon \geq 1$, and certain members of the generalized entropy class of inequality measures with $\alpha \leq 0$, which includes the commonly used mean log deviation.[38]

Sensitivity to very low incomes is, naturally, a desirable feature of distribution sensitive welfare measures. However, to prevent such measures from taking extreme values when there are a few very low values of y_i in the survey, it is standard practice to bottom-code income or consumption, depending on the survey, at some strictly positive value.[39] The practical issue is where to set the bottom-coding threshold. For surveys in which consumption is the main measure of well-being, very low reported consumption is likely to be the result of measurement error, particularly when reported values are lower than the biological minimum consumption level required to sustain life. For these surveys, we draw on estimates of the expenditure level required to obtain 2,330 kcal per day from the least expensive staple starches available in a country, as reported in Herforth et al. (2022). Across countries in 2017, the minimum cost of this bundle was \$0.24 per day in 2017 \$PPP. We round this to \$0.25 per day, and bottom-code consumption data at this threshold in our empirical analysis in Section 4.2.

For surveys in which income is the main measure of well-being, we cannot appeal to biological arguments to identify a bottom-coding threshold, since very low, zero, and even negative incomes are logically possible to observe when individuals can finance consumption by drawing down assets. Instead, we simply bottom-code income at the same threshold of \$0.25 per day, recognizing that the consumption levels of the bottom-coded individuals are unlikely to be lower given the biological estimates above, and could well be considerably higher.[40]

Of course, the estimates of the minimum level of expenditure needed to maintain life, on which we base our bottom-coding threshold, are, themselves, subject to measurement error. To explore the sensitivity of our results in Section 4.2, we consider a range of alternative values for the bottom-coding threshold. For the upper end of this range, we rely on estimates of the "consumption floor" proposed by Ravallion (2016). Implementing this methodology in 766 household surveys in PIP in which consumption is the main measure of well-being, we find that the lowest value across all

[38] The concern that distribution sensitive measures are highly sensitive to low incomes is not new. Cowell and Victoria-Feser (1996) discuss this issue in the context of data contamination (e.g., measurement error), using influence functions to demonstrate the extent to which a poverty or inequality measure is changed by small errors in the data.

[39] For example, the Luxembourg Income Study bottom- (and top-) codes the distribution of log income at three times the inter-quartile range below (above) the first (third) quartile when reporting inequality measures (Neugschwender, 2020). This means that the threshold at which incomes are truncated varies across surveys and will be lower in surveys in which the overall dispersion of measured incomes is high.

[40] To corroborate this, we examine 20 income surveys in the PIP database, in which measured incomes are zero in the lowest percentiles of the income distribution, *and* a consumption survey is available in the same year. These countries are all upper middle-income or high-income countries in Eastern Europe. Pooling across all surveys, median consumption in the bottom percentiles of the consumption distribution corresponding to the percentiles of the income distribution in which reported incomes are zero is \$2.64/day, and the lowest consumption percentile is \$0.84/day.

surveys is $0.48 per day (for Zambia in 1991), which we round to $0.50 per day.[41] We set the lower end of our range at $0.10 per day, corresponding to the median value of the bottom-coding threshold that we would obtain if we were to apply the LIS methodology (see footnote 25 above) in the 233 surveys carried out in the low-income countries in our dataset.

Figure A.1 illustrates the results of this sensitivity analysis. In Panel A, we show the trend in the global welfare gap, bottom-coding at $0.10, $0.25, and $0.50 per day. The choice of the bottom-coding threshold matters only minimally for the global welfare aggregate. In 1990, the global welfare gap ranges from 10.8 to 11.4 for the high and low ends of the range, as compared with 11.1 for our baseline bottom-coding threshold of $0.25 per day. Over time, the differences between the three measures become negligible, as growth in the past 30 years contributes to a much lower frequency of very low measured income or consumption in surveys.

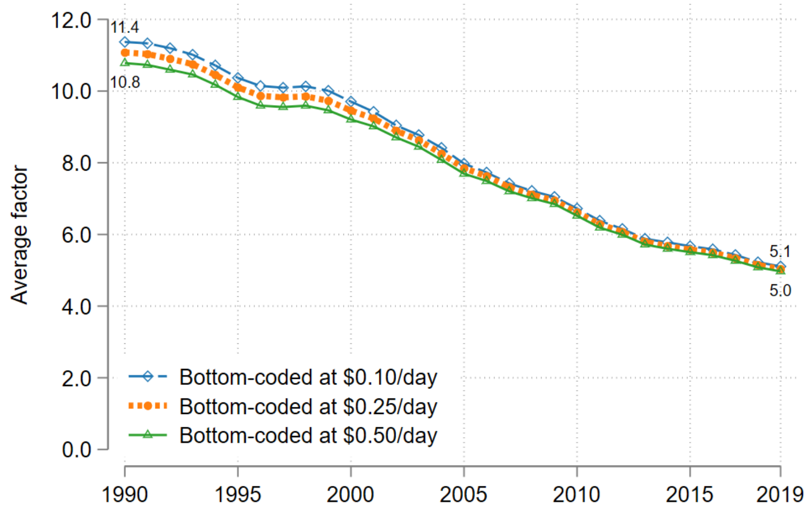
In Panels B and C, we compare country-level estimates of the welfare gap with different bottom-coding thresholds. In both panels, the baseline bottom-coding threshold of $0.25 per day is on the horizontal axis, and two alternative thresholds are on the vertical axes. Each dot represents a survey, and the color-coding distinguishes surveys before and after 2000. In both panels, the vast majority of surveys are very close to the 45-degree line, indicating that the choice of bottom-coding threshold matters little for the welfare gap estimates. The few exceptions are primarily very poor countries in Sub-Saharan Africa in the 1990s. In both panels, the Pearson and Spearman rank correlation coefficients between the baseline and alternative estimates are higher than 0.98.

Given the sensitivity of the welfare gap, as well as some other existing welfare and inequality measures, to very low – and likely mismeasured – consumption or income levels, some form of bottom-coding is needed in empirical applications to prevent a few observations with very low values from dominating the welfare gap. Based on biological considerations, we argue that $0.25 per day is a reasonable bottom-coding threshold and find that our empirical results vary only minimally over the range of plausible values around this threshold.

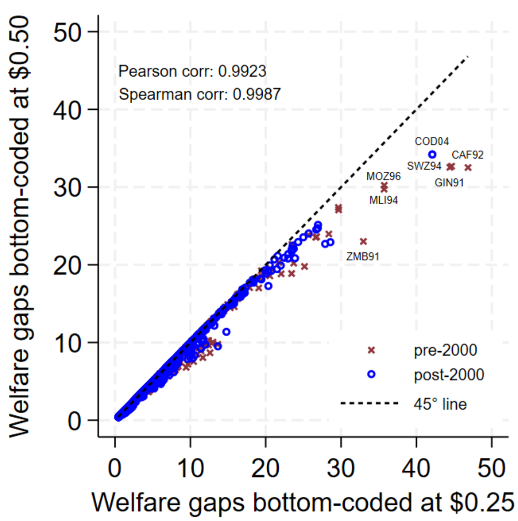
[41] Ravallion (2016) proposes a model of measurement error in consumption that can be smoothed out using a weighted sum of the observed consumption levels below the International Poverty Line ($2.15 per day in 2017 PPPs). The weights monotonically decline as consumption increases up to the poverty line.

Figure A.1: Sensitivity to very low welfare values

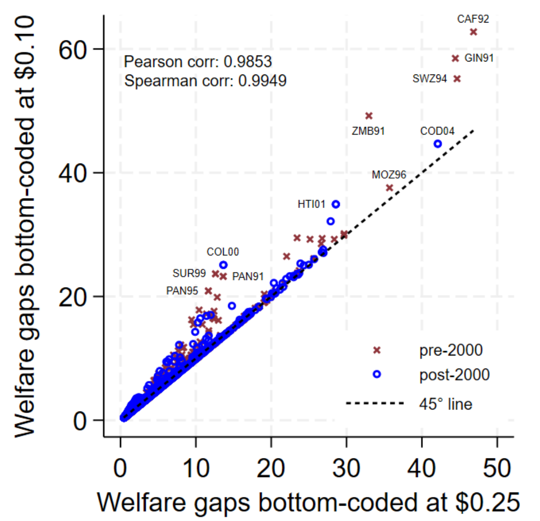
Panel A: Global welfare gap for different bottom-coding thresholds



Panel B: Welfare gap correlations, $0.25 or $0.50 per day



Panel C: Welfare gap correlations, $0.25 or $0.10 per day



Source: Authors' calculation using Poverty and Inequality Platform (September 2022).

Note: Panel A of this figure reports the trends in global welfare gap using $0.10, $0.25, and $0.50 per day as the bottom-coding threshold. Panel B reports the correlation of $W(y, z)$ bottom-coded at $0.25 or $0.50 per day and Panel C reports the correlation of $W(y, z)$ bottom-coded at $0.25 or $0.10 per day.