The Journal of Economic Inequality
https://doi.org/10.1007/s10888-025-09677-6

RESEARCH

Distributional divergence as a unifying measurement framework

Casilda Lasso de la Vega[1] · Juan Gabriel Rodríguez[2] · Rafael Salas[2]

Received: 20 November 2024 / Accepted: 22 April 2025
© The Author(s) 2025

Abstract
The measurement of various economic phenomena, including income inequality, poverty, richness, inequality-of-opportunity, horizontal equity, mobility, wage discrimination and goodness of fit, are examples of the measurement of the divergence between distributions. In all these examples there is a reference distribution and the aim is to examine to what extent any given distribution differs from the reference one. By providing an axiomatic characterization of a family of indices for measuring divergence between distributions, this paper provides a unifying framework for incorporating the measurement of many economic phenomena, while exploring the relationship of the derived family of indices to other existing measures.

Keywords Divergence · Characterization · Inequality · Poverty · Wage discrimination · Inequality of opportunity

JEL Classification D31 · D63 · I32 · J71

1 Introduction

The measurement of various economic phenomena often relies on indicators that assess the extent to which a given distribution deviates from a reference distribution. One prominent example is income inequality, where the actual income distribution is compared to a reference distribution that assigns the mean income to each individual. Other instances include measuring inequality of opportunity, where a "fair" distribution, free from circumstances, is used as a reference; poverty, where the distribution of the poor is compared to the poverty line; and wage discrimination, where the observed wage distribution is compared to an estimated wage distribution without discrimination. In all these cases, each individual in society is assigned

The original online version of this article was revised due to a retrospective Open Access order.

✉ Juan Gabriel Rodríguez
juangabr@ucm.es

Casilda Lasso de la Vega
casilda.lassodelavega@ehu.es

Rafael Salas
r.salas@ucm.es

[1] Dept. Quantitative Methods, University of the Basque Country (UPV/EHU), Biscay, Spain

[2] ICAE, EQUALITAS and CEDESOG, Universidad Complutense de Madrid, Madrid, Spain

Published online: 14 May 2025

two outcome values, and the indices gauge the differences between the corresponding distributions, making them specific examples of distributional divergence measurement.

This paper proposes a unifying framework to incorporate by means of divergence measurement some important economic phenomena such as those mentioned above and shows that this can be done by a small modification of some well-known properties, such as the Pigou-Dalton principle of transfers or the Decomposability a lá Shorrocks. This proposal can therefore help researchers and policy makers to adopt a broader perspective when measuring different economic phenomena such as inequality, poverty, richness, wage discrimination and inequality of opportunity. Along with traditional properties such as *continuity*, *anonymity*, *population invariance*, and *normalization* that are uncontroversial and widely shared among indices applied in empirical studies, this work incorporates novel elements that improve both theoretical and empirical applicability. The first innovation is the assumption of *decomposability* a lá Shorrocks, meaning that an increase in divergence within any subgroup leads to a corresponding increase in the overall population. This property is particularly advantageous for policy makers, as it allows disaggregation of divergence into population subgroups and enables targeted policy interventions to reduce divergence within specific population segments, while ensuring an overall reduction in distributional divergence. Empirical researchers also benefit from the possibility of closely examining and assessing the impact of economic growth and public policies on the different subgroups that make up society. This property is satisfied by broad families of measures, although it is a strict requirement that, among others, Gini-type measures fail to meet.

We also explore the implications of adopting the *generalized Pigou-Dalton principle* and the *gapping axiom*, both introduced by Almås et al. (2011) within the framework of inequality of opportunity. Building on the cornerstone property in the measurement of income inequality, the Pigou-Dalton principle - which states that a transfer of income from a richer to a poorer individual reduces inequality -, this study extends the idea to distributional divergence. The generalized Pigou-Dalton principle establishes that any progressive transfer from the gap of the richer to the gap of the poorer should reduce divergence. This generalization strengthens the theoretical foundation of divergence measures, and ensures that the measurement of divergence aligns with widely accepted redistribution principles. On the other hand, the gapping axiom stipulates that when the mean of actual outcomes equals the mean of reference outcomes, any sequence of changes that preserves the gap distribution and leaves the means unchanged should not modify the divergence level. This requirement improves the robustness of divergence indices since it introduces a crucial stability property that was missing from previous frameworks. To enable meaningful comparisons between societies with different total outcomes, the framework incorporates three additional properties: *scale invariance* which ensures that the divergence measure remains unchanged when all outcomes are multiplied by the same positive constant; *translation invariance* which guarantees that an equal increase in all outcomes does not alter the level of divergence; and *unit-consistency* (Zheng 2007a, b) which is a more relaxed version of scale invariance (only the ranking should remain unchanged when outcomes are multiplied by the same positive constant).

As shown below, a significant advantage of the proposed approach is that it extends the analysis of distributional divergence to several existing measures in different economic fields. Complementarily, the proposed framework allows obtaining new families of indices that may be useful to the applied researcher and policy maker interested in these fields.

In inequality measurement our framework encompasses well-known indices such as the coefficient of variation, the Robin Hood index, the Kolm-Pollack family, the variance, the family of absolute and relative compromises indices characterized by Ebert (1988), and a subfamily of unit consistent indices proposed in Zheng (2007a). In poverty analysis, it generalizes popular indices, including the $FGT-$family of relative and absolute poverty indices, the absolute poverty index proposed by Zheng (2007b), and the Krtscha-type intermediate poverty index. In richness metrics, it incorporates indices such as the Peichl-Schaefer-Schneider convex family and the richness index proposed by Medeiros (2006). In income mobility, the indices characterized by Mitra and Ok (1998) fit the setting, while in wage discrimination, our framework includes the convex subfamily of relative and absolute indices proposed by Jenkins (1994), and the family of relative and absolute indices proposed by del Río et al. (2011). As indices of goodness of fit (see also Cowell et al. (2015)), we find the mean square error (MSE) and the reduced Chi-square statistic in OLS. In addition, new indexes are obtained in all these fields. This fact is most evident in the measurement of inequality of opportunity where we obtained three new classes of inequality of opportunity indices - relative, absolute, and unit consistent - offering a more comprehensive set of tools for assessing fairness and unjust economic disparities.

There are other characterizations of divergence measures in the literature. The first axiomatization of functional forms of divergence measures is due to Cowell (1985). The main difference with our paper is that it assumes *differentiability*, two strong versions of *homotheticity* and a *monotonicity in distance* property stronger than the generalized Pigou-Dalton principle. More recently, Magdalou and Nock (2011), also assuming differentiability, characterizes the functional form of a family of measures of divergence that fulfills *judgment separability*, the key property in its characterization result. However, this property has no effect when we restrict attention to societies for which the means of the actual and reference distributions coincide. More importantly, none of the relative divergence measures derived in these two papers satisfies the gapping principle.

The remainder of the paper is organized as follows. Section 2 introduces the notation, axioms, and presents the characterization results. In Section 3, the derived families of divergence indices are compared with existing measures, and new indices are explored. Finally, Section 4 concludes the paper.

2 Measuring distributional divergence: axioms and characterization results

In this section we characterize families of distributional divergence measures. First, we present the basic definitions and the axioms needed for the characterizations.

2.1 Basic definitions

A population is a finite set $N \subset \mathbb{N}$, and the set of all populations is $\mathcal{N}$. Given a population $N_S \in \mathcal{N}$, a society is denoted by $S = (Y^A, Y^R)$, with $(Y^A, Y^R) \in \mathbb{R}^{N_S} \times \mathbb{R}^{N_S}$, where Y^A is the distribution of *actual* outcomes, and Y^R is the *reference* distribution. For each individual $i \in N_S$, $g_i = y_i^A - y_i^R$ represents the gap between the actual and the reference outcomes. We denote by $n(S)$ the population size and by $\mu^A(S)$ and $\mu^R(S)$ the outcome means of the actual and the reference distribution respectively. We assume that $n(S) \geq 3$ and that $\mu^A(S), \mu^R(S) > 0$. To simplify the notation we will sometimes write $\theta(S) = (n(S), \mu^A(S), \mu^R(S))$; and if there is no room for confusion, we will sometimes use the notation n, μ^A and μ^R to represent the population size and the respective means. Also, for the sake of clarity, we will occasionally denote societies as follows $S = \langle (y_1^A, y_1^R), (y_2^A, y_2^R), \ldots, (y_n^A, y_n^R) \rangle$. Given a society $S = (Y^A, Y^R)$, we denote by $\overline{S} = (X, Y^R)$, with $X = Y^R + (\mu^A(S) - \mu^R(S))$, the *equalized* version of S, that is, the society in which all the individual's gaps are equal, i.e., $g_i = \mu^A(S) - \mu^R(S)$ for all $i \in N_S$. The set of all societies is represented by $\mathbf{S}$ while $\mathbf{S}^*$ represents the subfamily of societies with $\mu^A(S) = \mu^R(S)$.

For any society $S = (Y^A, Y^R)$ and any scalar $\lambda > 0$, let $\lambda S = (\lambda Y^A, \lambda Y^R)$ denote the society obtained by multiplying all the outcomes by λ. Similarly, let $S+\lambda = (Y^A+\lambda, Y^R+\lambda)$ denote the society obtained by adding λ to all the outcomes. For any integer $k > 0$, let $S^{(k)}$ denote the society obtained by replicating k-times the outcomes of each of the individuals, that is, $n(S^{(k)}) = kn(S)$, and $S^{(k)} = (Y^{A(k)}, Y^{R(k)})$, with $Y^{A(k)} = (\underbrace{Y^A, Y^A, \ldots, Y^A}_{k-times})$ and $Y^{R(k)} = (\underbrace{Y^R, Y^R, \ldots, Y^R}_{k-times})$. For any two societies S and S', let $S \uplus S' = (S, S')$ denote the society obtained by concatenating societies S and S'.

A *distributional divergence index* is a function $J : \mathbf{S} \rightarrow \mathbb{R}_+$, which assigns to each society S a real number $J(S) \geq 0$, which is meant to capture the divergence between the actual and the reference distributions.

Note that when a distributional divergence index is restricted to societies where the reference distribution corresponds to the distribution of mean incomes, i.e., $Y^R = (\mu, \ldots, \mu)$ where $\mu = \mu^A = \mu^R$, it becomes an inequality measure.

2.2 Axioms

The following axioms will play a role in the characterizations. First of all *continuity*, which guarantees that similar societies have similar levels of distributional divergence.

Continuity $J(S)$ is a continuous function of the individual's actual and reference outcomes.

Permuting a society is only to change the order of the individuals, while (k)-replicating a society is to clone each individual k times. *Anonymity* and *population invariance* state, respectively, that the distributional divergence index J is not sensitive to these two transformations.

Anonymity For any two societies $S = (Y^A, Y^R)$ and $S' = (Y'^A, Y'^R)$, $J(S) = J(S')$ whenever $S' = (Y^A\pi, Y^R\pi)$ for some permutation matrix π.[1]

Population invariance For any society S and integer $k > 0$, $J(S) = J(S^{(k)})$.

Anonymity guarantees that no other attributes except the actual and the reference outcome of each individual matter in order to assess the level of divergence in a society. Population invariance requires that the level of divergence remains unaffected if we clone each individual a fixed number of times. What matters is not the absolute number of people who have any pair of actual and reference outcomes, but rather their proportion in the population. It enables the comparison of populations with different numbers of individuals.

Normalization is a cardinal property that demands that there is no divergence in the society if and only if the actual outcome and the reference outcome coincide for all the individuals.

Normalization For any society $S = (Y^A, Y^R)$, $J(S) = 0$ if and only if $y_i^A = y_i^R$ for all $i = 1, \ldots, n$.

A consequence of normalization is that if $J(S) = 0$, then $\mu^A(S) = \mu^R(S)$.

[1] A permutation matrix is a square matrix with only 0s and 1s, and such that the sum of each column and each row is equal to 1.

The next axiom, *decomposability*, is essentially what is known as subgroup consistency and is closely related to the notion of aggregativity of an index (see Shorrocks (1984, 1988)). It states that if a given society is composed of two regions, and the distributional divergence in one of its regions increases, then the distributional divergence in the whole society increases as well. The satisfaction of this axiom justifies the implementation of distributive policies in sub-regions in order to obtain results for the whole society. Despite its intuitive appeal, decomposability is a strong requirement that constrains the functional form of the measure. While, as we will show, it holds for broad families of measures, there are also widely used indices - such as Gini-type measures and indices based on income shares - that do not satisfy this property. The cardinal version of this axiom is formulated as follows.

Decomposability There exists an aggregator function A such that, for any two societies S and T, $J(S \uplus T) = A(J(S), \theta(S), J(T), \theta(T))$, where A is continuous and strictly increasing in $J(S)$ and $J(T)$.

The following properties allow us to compare societies with different total outcomes.

Scale invariance For any society S and scalar $\lambda > 0$, $J(S) = J(\lambda S)$.

Translation invariance For any society S and scalar $\lambda > 0$, $J(S) = J(S + \lambda)$.

Scale invariance requires that only the relative distribution of the respective outcomes determines the level of divergence. In other words, it is not necessary to know the units in which outcome is measured (dollars, euros, etc.) in each society to determine whether one has more or less distributional divergence than another. Measures that are scale invariant are referred to as *relative*. Relative measures can be expressed as a function of the outcomes normalized by the respective mean outcome. Translation invariance states that an increment by the same amount in all the outcomes does not affect the divergence level. Measures that satisfy translation invariance are referred to as *absolute*.

Either of these two axioms allows the comparison of populations with different total outcomes. However, these two axioms impose different value judgments in the measurement. Zheng (2007a, b) question the value judgments involved in the inequality and poverty relative measures and proposes, in the inequality and poverty fields, the unit-consistent axiom which requires that the *orderings* do not change under a change of units in which outcome is measured. The unit-consistency axiom can be introduced for a divergence measure as follows.

Unit consistency For any two societies S and S', if $J(S) \leq J(S')$, then $J(\lambda S) \leq J(\lambda S')$ for all $\lambda > 0$

Clearly, all relative measures are unit-consistent, but as we will show, they are not the only ones.

The above axioms represent standard properties that numerous economic indices satisfy. In addition to these general axioms, there are two specific ones that pertain to the measurement of distributional divergence. These two axioms outline the circumstances under which the divergence between distributions changes.

The subsequent axiom, known as the *generalized Pigou-Dalton principle*, plays a crucial role in determining the type of transformations that reduce distributional divergence. Introduced by Almås et al. (2011) as a weaker version of Cowell (1985)'s concept of *monotonicity in distance*,[2] this principle states that when two individuals in a society undergo an outcome transfer such that the individual with the higher gap gives up some of their actual outcome

[2] See expression (13) in Cowell (1985) taking $d(x, y) = x - y$. Moreover, the generalized Pigou-Dalton principle assumes that the distribution of reference outcomes remains unchanged.

to the other, the distributional divergence decreases. In other words, redistributive actions that close the gap between individuals reduce the overall divergence. This principle aligns with the standard Pigou-Dalton principle when the reference distribution assigns the mean outcome to each individual.

Generalized progressive transfer We say that society $S' = (Y'^A, Y^R)$ is obtained from society $S = (Y^A, Y^R)$ by a generalized progressive transfer if there are individuals j and k and $\delta > 0$ such that

- $g'_j = y'^A_j - y^R_j = g_j - \delta = y^A_j - \delta - y^R_j > g'_k = y'^A_k - y^R_k = g_k + \delta = y^A_k + \delta - y^R_k$
- $g'_l = y'^A_l - y^R_l = g_l = y^A_l - y^R_l$ for all $l \neq j, k$.

Note that if S' can be obtained from S by a generalized progressive transfer, then $\theta(S) = \theta(S')$.

Generalized Pigou-Dalton principle For any two societies S and S', if S' can be obtained from S by a generalized progressive Pigou-Dalton transfer, then $J(S') < J(S)$.

A straight consequence of the Generalized Pigou-Dalton principle is that if S' can be obtained from S by a sequence of progressive transfers then $J(S') < J(S)$. Accordingly, for any given society S, or $S = \overline{S}$, or $J(\overline{S}) < J(S)$.

Now we analyze more in detail the generalized Pigou-Dalton principle. From the above definition, a generalized progressive transfer occurs between two individuals j and k with $y^A_j - y^R_j > y^A_k - y^R_k$, that is, $y^A_j > y^A_k + (y^R_j - y^R_k)$. We find three possible cases.

i. If $y^R_j \geq y^R_k$, then $y^A_j > y^A_k + (y^R_j - y^R_k) \geq y^A_k$. The transfer occurs from a richer individual with a higher (or equal) reference outcome, to a poorer individual with a lower (or equal) reference outcome. In this case, the transfer is a standard progressive transfer from the richer to the poorer one.

ii. If $y^R_j < y^R_k$ and $y^A_j > y^A_k$. The transfer occurs from a richer individual with a lower reference outcome to a poorer individual with a higher reference outcome. Also in this case the transfer is a standard progressive transfer from the richer to the poorer one.

iii. If $y^R_j < y^R_k$ and $y^A_j \leq y^A_k$. The transfer occurs from the poorer individual j to the richer individual k because the reference outcome of the poorer individual is small enough so that the gap of j is greater than the gap of k. This is a regressive transfer. Thus, regressive transfers are possible when the poorer individual is better off than the richer one in terms of reference distribution (measured by the outcome gap).

To give the intuition of this third case, suppose that the reference distribution is given by the fair income of each individual which in turn is a direct function of the effort made.[3] Then, the transfer occurs from the poorer individual j to the richer individual k because j-individual's effort is low enough so that j-individual's gap is larger than k-individual's gap. When not only the actual outcome but also the reference outcome (e.g. fair income or effort) matters, there may be a trade-off between both dimensions, observed (income) and reference (effort) outcomes, so that regressive transfers in terms of outcomes but progressive in terms of gaps are allowed.

Finally, the last property, the *gapping axiom*, also introduced by Almås et al. (2011) as *unfairism* in its context, dictates a basic behavior for any measure of distributional divergence. This axiom pertains to societies $S \in \mathbf{S}^*$, where $\mu^A(S) = \mu^R(S)$. It requires that any sequence

[3] This example will gain full force in the next section, when we explain that inequality of opportunity is a particular case of distributional divergence.

of alterations in individual outcome values, provided it preserves the means and leaves the gap distribution unchanged, should not affect the level of divergence. To express it formally, the gapping axiom is defined as follows.

Gapping axiom For any two societies $S, S' \in \mathbf{S}^*$, with the same population and the same actual and reference outcome means, that is, $\theta(S) = \theta(S')$, if $g_i = g_i'$ for all $i = 1, \ldots, n$, then $J(S) = J(S')$.

One may wonder the implications of redefining the gapping axiom to consider relative gaps. Given that both the actual and reference distributions in the two involved societies have identical means, the gapping axiom can be equivalently expressed in terms of relative gaps divided by the corresponding means without altering its fundamental concept.

It is important to emphasize that, as shown in the upcoming section, this axiom is not met by the class of relative measures of divergence characterized in both Cowell (1985) and Magdalou and Nock (2011).

2.3 Characterization results

We now present the results. All the proofs are in the Appendix.

The first theorem shows that the divergence indices we are looking for have an additive structure. This is a standard result when *decomposability* is assumed.

Theorem 1 *A continuous distributional divergence index J satisfies anonymity, population invariance, normalization, decomposability and the generalized Pigou-Dalton principle if and only if there exist functions F and ϕ such that for any society $S \in \boldsymbol{S}$*

$$F\left(J(S), \mu^A(S), \mu^R(S)\right) = \frac{1}{n(S)} \sum_{i=1}^{n(S)} \left\{\phi(y_i^A, y_i^R) - \phi(\mu^A(S), \mu^R(S))\right\}, \qquad (1)$$

where $F : D \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous in its three components, strictly increasing in the first one, and satisfies that, for any $(x, y, z) \in D$ with $F(x, y, z) = 0$, then $x = 0$ if and only if $y = z$.
In addition, $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a continuous function, with $\phi(x, x) = ax + b$, and such that $\phi(x + \delta, y) + \phi(x' - \delta, y') > \phi(x, y) + \phi(x', y')$ for all $\delta > 0$, whenever $x - y \geq x' - y'$.

Proof See Section 5.1 in the Appendix. □

Theorem 1 characterizes a broad family of divergence indices, encompassing many well-known measures from the literature. However, certain indices - such as Gini-type measures and those based on income shares - are not decomposable and therefore fall outside this family. Note that the functional form identified in Eq. 1 involves a function F that, for each society S, depends not only on its index value, $J(S)$, but also on its actual and reference means, $\mu^A(S)$ and $\mu^R(S)$. This implies that, given a function ϕ we are able to compare all the societies with same actual and reference means, and societies with different means are compared through different functions F. Indeed, Eq. 1 can be rewritten as

$$J(S) = F^{-1}\left(\frac{1}{n(S)} \sum_{i=1}^{n(S)} \left\{\phi(y_i^A, y_i^R) - \phi(\mu^A(S), \mu^R(S))\right\}, \mu^A(S), \mu^R(S)\right),$$

where the inversion of F is taken with respect to its first argument.

We briefly mention two families of well-known divergence indices that fit Eq. 1.

- Focusing on societies in which both the actual and the reference incomes are all positive, Csiszár (1991) characterizes the Bregman-Csiszár's family, whose symmetric version, using the notations of this paper, is given by $H_\alpha(S) = \sum_{i=1}^{n(S)} h_\alpha(y_i^A, y_i^R)$, where $\alpha \leq 1$ and h_α is as follows

$$h_\alpha(x, y) = \begin{cases} \frac{1}{\alpha}(y^\alpha - x^\alpha) + y^{\alpha-1}(x - y) & \text{if } \alpha < 1 \quad \text{and } \alpha \neq 0, \\ ln\frac{y}{x} + \frac{y}{x} - 1 & \text{if } \alpha = 0, \\ xln\frac{x}{y} - x + y & \text{if } \alpha = 1. \end{cases}$$

For each $\alpha \leq 1$, we consider $F_\alpha\left(H_\alpha(S), \mu^A, \mu^R, n\right) = \frac{1}{n}\left(H_\alpha(S) - h_\alpha(\mu^A, \mu^R)\right)$. The resulting index fits Eq. 1 with $\phi_\alpha = h_\alpha$.

- The family of relative divergence measures characterized by Cowell (1985), still focusing on societies with positive incomes, can be rewritten, using the notation of this paper, as follows

$$J_\alpha(S) = \begin{cases} \frac{1}{\alpha(\alpha-1)n} \sum_{i=1}^{n} \left(\left(\frac{y_i^A}{\mu^A}\right)^\alpha \left(\frac{y_i^R}{\mu^R}\right)^{1-\alpha} - 1\right) & \text{if } \alpha \notin \{0, 1\}, \\ \frac{1}{n} \sum_{i=1}^{n} \frac{y_i^R}{\mu^R} \; ln\left(\frac{y_i^R/\mu^R}{y_i^A/\mu^A}\right) & \text{if } \alpha = 0, \\ \frac{1}{n} \sum_{i=1}^{n} \frac{y_i^A}{\mu^A} \; ln\left(\frac{y_i^A/\mu^A}{y_i^R/\mu^R}\right) & \text{if } \alpha = 1. \end{cases}$$

Using the transformation $F_\alpha(J_\alpha(S), \mu^A, \mu^R, n) = J_\alpha(S)(\mu^A)^\alpha(\mu^R)^{1-\alpha}$, we find that the resulting family of indices fits Eq. 1 with

$$\phi_\alpha(x, y) = \begin{cases} \frac{x^\alpha y^{1-\alpha}}{\alpha(\alpha-1)} & \text{if } \alpha \notin \{0, 1\}, \\ yln\frac{y}{x} & \text{if } \alpha = 0, \\ xln\frac{x}{y} & \text{if } \alpha = 1. \end{cases}$$

None of the above examples satisfies the gapping axiom.[4] The consequences of assuming this property are explored in the following theorem.

Theorem 2 *A continuous distributional divergence index J satisfies anonymity, population invariance, normalization, decomposability, the generalized Pigou-Dalton principle and the gapping axiom if and only if there exist functions F and φ such that for any $S \in \boldsymbol{S}$*

$$F\left(J(S), \mu^A(S), \mu^R(S)\right) = \frac{1}{n(S)} \sum_{i=1}^{n(S)} \left\{\varphi(y_i^A - y_i^R) - \varphi(\mu^A(S) - \mu^R(S))\right\}, \tag{2}$$

where $F : D \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous in its three components, strictly increasing in the first one, and satisfies that, for any $(x, y, z) \in D$ with $F(x, y, z) = 0$, then $x = 0$ if and only if $y = z$. In addition, $\varphi : \mathbb{R} \to \mathbb{R}$ is a continuous strictly convex function.

Proof See Section 5.2 in the Appendix. □

The results obtained are still very general. Now we examine the consequences of imposing some invariance conditions.

[4] To show this, consider two societies $S = \langle(1, 2), (9, 6), (2, 4)\rangle$ and $S' = \langle(3, 4), (6, 3), (3, 5)\rangle$ with $n_S = n_{S'} = 3$, $\mu^A(S) = \mu^R(S) = \mu^A(S') = \mu^R(S') = 4$, and $g_1 = g_1' = -1$, $g_2 = g_2' = 3$ and $g_3 = g_3' = -2$. Then, according to the gapping axiom, the values of the indices for these two societies should be the same. However, upon closer examination, it becomes evident that this is not the case.

Theorem 3 *A continuous distributional divergence index J satisfies anonymity, population invariance, normalization, decomposability, the generalized Pigou-Dalton principle, the gapping axiom and scale invariance if and only if there exist a function F and a parameter $\alpha > 1$ such that for any $S \in \boldsymbol{S}$*

$$F\left(J(S), \mu^A(S), \mu^R(S)\right) = \frac{1}{n(S)} \sum_{i=1}^{n(S)} K_i \left\{\left|y_i^A - y_i^R\right|^{\alpha} - \left|\mu^A(S) - \mu^R(S)\right|^{\alpha}\right\}, \qquad (3)$$

where $K_i = \frac{A^1+A^2}{2} + \frac{(A^1-A^2)sgn(y_i^A-y_i^R)}{2}$ with $A^1, A^2 > 0$.[5] *In addition, $F : D \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous in its three components, strictly increasing in the first one, and satisfies that, for any $(x, y, z) \in D$ with $F(x, y, z) = 0$, then $x = 0$ if and only if $y = z$. F also satisfies that $F(., \lambda x, \lambda y) = \lambda^{\alpha} F(., x, y)$, for all $x, y, \lambda > 0$.*

Moreover, for any $S \in \boldsymbol{S}^*$

$$\Phi(J(S)) = \frac{1}{n(S)} \sum_{i=1}^{n(S)} K_i \frac{\left|y_i^A - y_i^R\right|^{\alpha}}{\mu^{\alpha}}, \qquad (4)$$

where Φ is a continuous and strictly increasing function, and $\mu = \mu^A(S) = \mu^R(S)$.

Proof See Section 5.3 in the Appendix. □

Remark 1 The theorem above characterizes divergence indices that exhibit scale invariance and identifies how they quantify divergence in societies where the total actual and reference outcomes are identical. It is evident that a single extension of these indices to encompass all other types of societies does not exist. Below, we present two illustrative examples.

i.

$$J(S) = \frac{1}{n} \sum_{i=1}^{n} K_i \frac{\left|y_i^A - y_i^R\right|^{\alpha} - \left|\mu^A - \mu^R\right|^{\alpha}}{(\gamma\mu^A + (1-\gamma)\mu^R)^{\alpha}} \text{ for all } S \in \mathbf{S}, \text{ with } \alpha, \gamma \in \mathbb{R}, \alpha > 1 \text{ and } \gamma \in [0, 1].$$

ii.

$$J(S) = \frac{1}{n} \sum_{i=1}^{n} K_i \frac{\left|y_i^A - y_i^R\right|^{\alpha} - \left|\mu^A - \mu^R\right|^{\alpha}}{(\mu^A)^{\gamma}(\mu^R)^{\alpha-\gamma}} \text{ for all } S \in \mathbf{S}, \text{ with } \alpha, \beta, \gamma \in \mathbb{R}, \alpha > 1 \text{ and } \gamma \in [0, 1].$$

Note that any increasing transformation of the relative indices showed above collapses into Eq. 4 when $\mu^A = \mu^R$.

The following result offers additional clarification regarding the absolute indices.

Theorem 4 *A continuous distributional divergence index J satisfies anonymity, population invariance, normalization, decomposability, the generalized Pigou-Dalton principle and translation invariance if and only if there exist functions F and φ such that for any $S \in \boldsymbol{S}$*

$$F\left(J(S), \mu^A(S), \mu^R(S)\right) = \frac{1}{n(S)} \sum_{i=1}^{n(S)} \left\{\varphi(y_i^A - y_i^R) - \varphi(\mu^A(S) - \mu^R(S))\right\}, \qquad (5)$$

[5] Or, equivalently,

$$K_i = \begin{cases} A_1 & \text{if } y_i^A \geq y_i^R \\ A_2 & \text{if } y_i^A < y_i^R \end{cases}$$

where $F : D \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous in its three components, strictly increasing in the first one, and satisfies that, for any $(x, y, z) \in D$ with $F(x, y, z) = 0$, then $x = 0$ if and only if $y = z$. In addition, $F(., x + \lambda, y + \lambda) = F(., x, y)$, for all $x, y, \lambda > 0$, and $\varphi : \mathbb{R} \to \mathbb{R}$ is a continuous strictly convex function.

Proof See Section 5.4 in the Appendix. □

Note that Eq. 5 is equivalent to Eq. 2. As a result, Theorem 4 basically shows that all the divergent measures that satisfy the Pigou-Dalton principle fulfill the gapping axiom.[6] Finally we examine the implications of assuming unit-consistency.

Theorem 5 *A continuous distributional divergence index J satisfies anonymity, population invariance, normalization, decomposability, the generalized Pigou-Dalton principle, the gapping axiom and unit-consistency if and only if there exist a function F and a parameter $\alpha > 1$ such that for any $S \in \boldsymbol{S}$*

$$F\left(J(S), \mu^A(S), \mu^R(S)\right) = \frac{1}{n(S)} \sum_{i=1}^{n(S)} K_i \left\{ \left| y_i^A - y_i^R \right|^\alpha - \left| \mu^A(S) - \mu^R(S) \right|^\alpha \right\}, \tag{6}$$

where $K_i = \frac{A^1+A^2}{2} + \frac{(A^1-A^2)sgn\left(y_i^A - y_i^R\right)}{2}$ with $A^1, A^2 > 0$,[7] In addition, $F : D \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous in its three components, strictly increasing in the first one, and satisfies that, for any $(x, y, z) \in D$ with $F(x, y, z) = 0$, then $x = 0$ if and only if $y = z$. F also satisfies that $F\left(J(\lambda S), \mu^A(\lambda S), \mu^R(\lambda S)\right) = \lambda^\tau F\left(J(S), \mu^A(S), \mu^R(S)\right)$, for all $S \in \boldsymbol{S}$ and for all $\lambda > 0$, with $\tau \in \mathbb{R}$.

Moreover, for any $S \in \boldsymbol{S}^*$

$$\Phi(J(S)) = \frac{1}{n(S)} \sum_{i=1}^{n(S)} K_i \frac{\left|y_i^A - y_i^R\right|^\alpha}{\mu^\beta}, \tag{7}$$

where Φ is a continuous and strictly increasing function, $\mu = \mu^A(S) = \mu^R(S)$ and $\beta = \alpha + \tau$.

Proof See Section 5.5 in the Appendix. □

Remark 2 Note that when $\beta = \alpha$ the subfamily of relative indices is obtained. Taking into account the families given in Remark 1 it is straightforward to generate different families of unit-consistent divergence indices.

i.

$$J(S) = \frac{1}{n} \sum_{i=1}^{n} K_i \frac{\left|y_i^A - y_i^R\right|^\alpha - \left|\mu^A - \mu^R\right|^\alpha}{(\gamma\mu^A + (1-\gamma)\mu^R)^\beta} \text{ for all } S \in \mathbf{S}, \text{ with } \alpha, \beta, \gamma \in \mathbb{R}, \alpha > 1 \text{ and } \gamma \in [0, 1].$$

ii.

$$J(S) = \frac{1}{n} \sum_{i=1}^{n} K_i \frac{\left|y_i^A - y_i^R\right|^\alpha - \left|\mu^A - \mu^R\right|^\alpha}{(\mu^A)^\gamma (\mu^R)^{\beta-\gamma}} \text{ for all } S \in \mathbf{S}, \text{ with } \alpha, \beta, \gamma \in \mathbb{R}, \alpha > 1 \text{ and } \gamma \in [0, 1].$$

[6] Cowell (1985) obtains a similar result using differentiability.

[7] See footnote 5.

3 Examples of existing indices as particular cases of divergence and new indices derived

A measure of divergence is a function that assesses the disparity between the observed outcome distribution and a hypothetical scenario in which individuals receive outcomes corresponding to their reference values. In this section, we demonstrate that certain indices identified previously are widely recognized measures of income inequality, poverty, richness, inequality of opportunity, income mobility, (gender) wage discrimination, goodness of fit, and household differences in needs. Moreover, the divergence framework we have established enables the formulation of novel indices.

3.1 Income inequality

Income inequality indices evaluate the divergence between the actual distribution Y^A and the hypothetical scenario in which all individuals receive the mean outcome, that is, $Y^R = (\mu, \dots, \mu)$. Note that, by definition, $\mu^A = \mu^R$. Under this distribution of reference, we find that some existing relative, absolute and unit-consistent inequality indices belong to the families we have identified.

3.1.1 Relative inequality indices

We consider Eq. 4. We get the following results.

- When $\alpha = 2$ and $A_1 = A_2 = 1/2$, Eq. 4 gives the member of the Generalized Entropy family corresponding to the value of the parameter $\alpha = 2$, $GE_2(S) = \frac{1}{2n}\sum_{i=1}^{n}\frac{(y_i^A-\mu)^2}{\mu^2}$, which is ordinally equivalent to the coefficient of variation CV, i.e., $GE_2(S) = \frac{CV^2}{2}$.
- The *relative compromise indices* characterized by Ebert (1988) can be written, up to a positive (multiplicative) constant, as $\widetilde{E}_\alpha(S) = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{|y_i^A-\mu|^\alpha}{\mu^\alpha}\right)^{1/\alpha}$, with $\alpha > 1$. This family is an increasing transformation of the indices obtained from Eq. 4 with $A_1 = A_2 = 1$.
- The *Schutz coefficient* (or *Robin Hood index*) corresponds to the extreme case when the α-parameter tends to 1. In fact, $lim_{\alpha\to 1}\sum_{i=1}^{n}\left|y_i^A-\mu\right|^\alpha/2n\mu^\alpha = \sum_{i=1}^{n}|y_i^A-\mu|/2n\mu$.

3.1.2 Absolute inequality indices

Now we analyze the absolute indices that can be obtained from Eq. 5 through different specifications of φ.

- When $\varphi(x) = x^2$, from Eq. 5 we get the *variance*.
- The *absolute compromise indices* characterized by Ebert (1988) can be written, up to a positive (multiplicative) constant, as $E_\alpha(S) = \left(\frac{1}{n}\sum_{i=1}^{n}\left|y_i^A-\mu\right|^\alpha\right)^{1/\alpha}$, with $\alpha > 1$. This family is an increasing transformation of the class obtained in Eq. 5 taking $\varphi(x) = |x|^\alpha$.
- The *Chakravarty-Tyagarupananda* family of absolute indices (Chakravarty and Tyagarupananda 1998) is given by $CT_c(S) = \frac{1}{n}\sum_{i=1}^{n}\left(e^{c(y_i^A-\mu)}-1\right)$, with $c > 0$. This family fits Eq. 5 taking $\varphi(x) = e^{cx} - 1$.

- The *Kolm-Pollack family* of absolute indices (see Kolm (1976)) is given by $KP_\lambda(S) = \frac{1}{\lambda} ln \frac{1}{n} \sum_{i=1}^{n} e^{\lambda(\mu - y_i^A)}$, with $\lambda > 0$. It corresponds, up to multiplication by $1/\lambda$, to the specification $\varphi(x) = e^{-\lambda x}$ in Eq. 5.

3.1.3 Unit-consistent inequality indices

Regarding the unit-consistent indices, taking $\alpha = 2$ and $K = 1$ in Eq. 7 yields the members of the family characterized by Zheng (2007a) that satisfy the gapping axiom. This subfamily includes, as particular cases when $0 < \beta < 2$, the *generalization of the Krstcha intermediate* indices. In particular, the Krstcha index (Krtscha 1994) corresponds to the value $\beta = 1$.

3.2 Poverty

Following the work of Sen (1976), the measurement of poverty involves a two-step process. First, individuals living in poverty are identified using a poverty threshold, denoted as z. Then, their incomes are aggregated into a poverty indicator. As a result, any poverty measure satisfies the focus axiom, which establishes that only the incomes of those living in poverty are relevant for poverty measurement. Consequently, most poverty indices can be expressed in terms of the *(upper) censored income distribution*, introduced by Hamada and Takayama (1977). Given an income distribution Y^A and a poverty line z, the (upper) censored income distribution associated to Y^A and z, denoted by Y^{A^*}, is the distribution that sets all incomes above the poverty line to the poverty line. Formally, $y_i^{A^*} = min\left\{y_i^A, z\right\}$. Given a poverty line z, and an income distribution Y^A, a poverty index can be interpreted as the divergence between the censored income distribution Y^{A^*} and the reference distribution $Y^R = (z, z, ..., z)$. It's important to note that $\mu^{A^*} \leq \mu^R = z$. As demonstrated below, some of the most widely-used poverty indices can be regarded as specific cases of the family of divergence indices derived in Eq. 3.

- The FGT-family of poverty indices (Foster et al. 1984) is given by $FGT_\alpha(S) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z-y_i^{A^*}}{z}\right)^\alpha$, with $\alpha > 0$. Using the transformation $F(x_1, x_2, x_3) = x_1 x_3^\alpha - (x_3 - x_2)^\alpha$, we see that any index of the FGT-family with $\alpha > 1$, can be transformed into a member of the family identified in Eq. 3. Indeed,

$$F(FGT_\alpha(S), \mu^{A^*}, z) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z - y_i^{A^*}}{z}\right)^\alpha z^\alpha - (z - \mu^{A^*})^\alpha = \frac{1}{n}\sum_{i=1}^{n}\left|y_i^{A^*} - z\right|^\alpha - \left|\mu^{A^*} - z\right|^\alpha.$$

- Similarly, it is straightforward to show that both the absolute counterpart of the FGT-family given by $AFGT_\alpha(S) = \frac{1}{n}\sum_{i=1}^{n}\left(z - y_i^{A^*}\right)^\alpha$ (Atkinson 1992), and its unit-consistent version $IFGT_\alpha(S) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z-y_i^{A^*}}{1-\epsilon+\epsilon z}\right)^\alpha$, with $\epsilon \in (0, 1)$ (Zheng 2007b), also fit Eq. 2.
- The absolute poverty index proposed by Zheng (2000) can be written as follows: $P_\alpha(S) = \frac{1}{n}\sum_{i=1}^{n}(e^{\alpha(z-y_i^{A^*})} - 1)$, with $\alpha > 0$. Using the transformation $F(x_1, x_2, x_3) = x_1 - (e^{\alpha(x_3 - x_2)} - 1)$, we see that any P_α index can be transformed into a member of the family identified in Eq. 2 with $\varphi(x) = e^{\alpha|x_3 - x_2|} - 1$. Indeed,

$$F(P_\alpha(S), \mu^{A^*}, z) = \frac{1}{n}\sum_{i=1}^{n}\left(e^{\alpha(z-y_i^{A^*})} - 1\right) - \left(e^{\alpha(z-\mu^{A^*})} - 1\right) = \frac{1}{n}\sum_{i=1}^{n}\varphi(y_i^{A^*} - z) - \varphi(\mu^{A^*} - z).$$

- The Krtscha-type intermediate poverty index derived by Zheng (2007b) is as follows: $K_\beta(S) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z-y_i^{A^*}}{z^\beta}\right)^2$, with $0 < \beta < 1$. Using the transformation $F(x_1, x_2, x_3) = x_1 x_3^{2\beta} - (x_3 - x_2)^2$, we find that K_β is one of the members given in Eq. 2 with $\alpha = 2$. Indeed,

$$F(K_\beta(S), \mu^{A^*}, z) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z-y_i^{A^*}}{z^\beta}\right)^2 z^{2\beta} - (z-\mu^{A^*})^2 = \frac{1}{n}\sum_{i=1}^{n}\varphi(y_i^{A^*}-z) - \varphi(\mu^{A^*}-z).$$

3.3 Richness

The measurement of richness can be considered as the counterpart of the measurement of poverty, when the focus is on the rich people, who are identified through a richness line, ρ. People whose income is below ρ do not matter in the measurement of richness. Thus, any richness index can be defined in terms of the *(lower) censored income distribution*. Given an income distribution Y^A and a richness line ρ, the (lower) censored income distribution associated to Y^A and ρ, denoted by $Y^{A^{**}}$, is the distribution that sets all incomes below the richness line to the richness line. Formally, $y_i^{A^{**}} = max\left\{y_i^A, \rho\right\}$. Given a a richness line ρ, and an income distribution Y^A, a richness index can be interpreted as the divergence between the censored income distribution $Y^{A^{**}}$ and the reference distribution $Y^R = (\rho, \ldots, \rho)$. Note that $\mu^{A^{**}} \geq \mu^R = \rho$. We now show that some of the existing richness indices introduced in the literature can be considered as particular cases of the family derived in Eq. 2.

- The *Peichl-Schaefer-Schneider (convex)* family (Peichl et al. 2010) is the counterpart of the FGT family in the measurement of richness. The family can be written as $PSS_\alpha(S) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i^{A^{**}}-\rho}{\rho}\right)^\alpha$, with $\alpha > 1$. The same procedure followed in the case of the FGT family shows that this family is part of the indices identified in Eq. 3.
- The absolute version of the PPS-family is given by $M_\alpha(S) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i^{A^{**}} - \rho\right)^\alpha$.[8] Using the transformation $F(x_1, x_2, x_3) = x_1 - (x_2 - x_3)^\alpha$, we see that any index M_α with $\alpha > 1$, can be transformed into a member of the family identified in Eq. 2 with $\varphi(x) = x^\alpha$. Indeed,

$$F(M_\alpha(S), \mu^{A^{**}}, \rho) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i^{A^{**}} - \rho\right)^\alpha - (\mu^{A^{**}} - \rho)^\alpha = \frac{1}{n}\sum_{i=1}^{n}\varphi(y_i^{A^{**}} - \rho) - \varphi(\mu^{A^{**}} - \rho).$$

3.4 Inequality of opportunity

The measurement of inequality of opportunity differs from classical inequality analysis by considering individual effort alongside circumstances.[9] In this field, it is assumed that outcomes are influenced by factors beyond individuals' control (circumstances) and factors for which individuals are considered responsible (effort). Inequality resulting from circumstances is deemed ethically unacceptable, while that arising from effort is considered fair. The ultimate objective is to attain a distribution where the influence of circumstances is elim-

[8] When $\alpha = 1$, it corresponds to the richness index introduced by Medeiros (2006)'s.

[9] For comprehensive surveys, see, for instance, Pignataro (2012), Roemer and Trannoy (2016), and Ramos and van de Gaer (2016).

inated, leaving only inequality arising from effort. Three main approaches for measuring inequality of opportunity have been proposed.

First, direct measures exclusively consider inequality due to circumstances. Second, indirect measures gauge the remaining inequality after equalizing opportunities. Lastly, norm-based measures assess the disparity between an individual's actual outcome and a *circumstances-free fair* outcome. Our proposal falls within the category of norm-based measures.

In the literature, two main norm-based measures have been introduced. Devooght (2008) adopted the *egalitarian equivalent allocation* (Bossert and Fleurbaey 1996) as the norm and suggested using Cowell (1985)'s measure of distributional change. Alternatively, Almås et al. (2011) employed the *generalized proportionality allocation* (Bossert 1995) as the norm and proposed the unfairness Gini coefficient.

In line with this literature, the framework presented in this paper allows for the adoption of various allocations as norms, such as the *observable average egalitarian equivalent allocation* (Bossert et al. 1999) or the *conditional egalitarian mechanism* (Kolm 2002; Bossert et al. 1999). The key distinction from previous proposals lies in the wide range of divergence measures obtained, which can be applied to derive new indices of inequality of opportunity, whether relative, absolute, or unit-consistent.

Specifically, let $Y^F = \left(y_1^F, \ldots, y_n^F\right)$ represent a *fair* distribution in a given society, taken as the reference distribution. Assuming $\mu^A = \mu^R$, and based on Theorem 2, the following families of inequality of opportunity indices, all satisfying decomposability, the generalized Pigou-Dalton principle and the gapping axiom, can be derived. Depending on the equation chosen as a starting point, Eqs. 4, 5, or 7, the indices can be classified as relative, absolute, or unit-consistent, respectively. Among other possibilities, we obtain the following.

- $IOp_\alpha^R(S) = \frac{1}{n}\sum_{i=1}^n \frac{\left|y_i^A - y_i^F\right|^\alpha}{\mu^\alpha}$ with $\alpha > 1$, obtained from Eq. 4, is a family of relative inequality of opportunity measures.
- $IOp_\alpha^A(S) = \frac{1}{n}\sum_{i=1}^n (e^{\alpha(y_i^A - y_i^F)} - 1)$ with $\alpha > 0$, obtained from Eq. 5, is a family of absolute inequality of opportunity measures.
- $IOp_\alpha^U(S) = \frac{1}{n}\sum_{i=1}^n \frac{\left|y_i^A - y_i^F\right|^\alpha}{\mu^\beta}$ with $\alpha > 1$, obtained from Eq. 7, is a family of unit-consistent inequality of opportunity measures.

3.5 Income mobility

Income mobility pertains to the changes in the income distribution within a specific society over a defined period. The concept involves the comparison of a reference distribution, typically the distribution at time $t = 0$, with the distribution observed after a specific duration of time has elapsed.

- The family of absolute income mobility measures characterized by Mitra and Ok (1998) can be written, up to an increasing transformation, as $M(S) = \frac{1}{n}\sum_{i=1}^n \left|y_i^A - y_i^R\right|^\alpha$, with $\alpha \geq 1$. It is clear that any member of this family for $\alpha > 1$ can be obtained from Eq. 5.

3.6 Gender wage discrimination

Gender wage discrimination can be interpreted as the divergence between the female wage distribution, denoted by $Y^A = (w_{f1}^A, \ldots, w_{fn^*}^A)$, where n^* is the number of women in population, and the female wage distribution without discrimination denoted by $Y_f^R =$

$(r_{f1}, \ldots, r_{fn^*})$, taking into account that, in general, the means of these two distributions do not coincide. As shown below, a number of gender wage discrimination measures introduced in the literature fit Eq. 2.

- A sub-family of the absolute wage discrimination indices proposed by Jenkins (1994) (equation 12) can be written as $AJ_\alpha(S) = \frac{1}{\alpha n^*}\sum_{i=1}^{n^*}\left|r_{fi} - w_{fi}^A\right|^\alpha$, with $\alpha > 1$. Then, using the transformation $F(x_1, x_2, x_3) = \alpha x_1 - |x_3 - x_2|^\alpha$, we see that any index AJ_α with $\alpha > 1$, can be transformed into a member of the family identified in equation (2) with $\varphi(x) = |x|^\alpha$.
- The corresponding sub-family of relative indices (Jenkins 1994) is written as $RJ_\alpha(S) = \frac{1}{\alpha n^*}\sum_{i=1}^{n^*}\frac{\left|r_{fi}-w_{fi}^A\right|^\alpha}{(\mu^R)^\alpha}$, with $\alpha > 1$. In this case, using the transformation $F(x_1, x_2, x_3) = \alpha x_1 x_3^\alpha - |x_3 - x_2|^\alpha$, any index RJ_α with $\alpha > 1$, can be transformed into a member of the family identified in Eq. 3.

Jenkins' families of wage discrimination indices depend on the wage differentials of both discriminated and non-discriminated women. Alternatively, del Río et al. (2011) proposes to focus only on the wage gaps of discriminated women. Given wage distributions Y^A and Y^R, we can use the upper censored wage distribution associated to Y^A and Y^R, denoted by Y^{A^*}, defined as $w_{fi}^{A^*} = min\left\{w_{fi}^A, r_{fi}\right\}$.

- The *ADGC* family of absolute wage discrimination indices proposed by del Río et al. (2011) can be written as $ADGC_\alpha(S) = \frac{1}{n^*}\sum_{i=1}^{n^*}\left(r_{fi} - w_{fi}^{A^*}\right)^\alpha$, with $\alpha > 1$.
- The *RDGC* family of relative wage discrimination indices proposed by del Río et al. (2011) can be written as $RDGC_\alpha(S) = \frac{1}{n^*}\sum_{i=1}^{n^*}\frac{\left(r_{fi}-w_{fi}^{A^*}\right)^\alpha}{(\mu^R)^\alpha}$, with $\alpha > 1$.

It is straightforward to show that these two families can be obtained from Eq. 2 with $\varphi(x) = |x|^\alpha$ if we use the transformations $F(x_1, x_2, x_3) = x_1 - (x_3 - x_2)^\alpha$ and $F(x_1, x_2, x_3) = \alpha x_1 x_3^\alpha - (x_3 - x_2)^\alpha$, respectively.

To replicate some results from the literature on wage discrimination based on the distributional approach we have focused on women. However, following Prieto-Rodríguez et al. (2020) we could have considered a more general framework in which both women and men might be discriminated against. In this case, we could replicate the previous results, but for the distribution of (discriminated) workers rather than for the distribution of (discriminated) women.

3.7 Goodness of fit

Let $Y^A = (y_1^A, \ldots, y_n^A)$ be the vector of observed values of a variable we want to predict and $Y^R = (\hat{y}_1, \ldots, \hat{y}_n)$ be the vector of predicted values. Assume -as in the OLS estimation- that the mean of the two distributions Y^A and Y^R is the same. Then, the divergence indices identified in Eq. 5 can be used as a measure of goodness-of-fit of a statistical model (see Cowell et al. (2015)). In fact, we get the following results.

- Taking $\varphi(x) = x^2$ in Eq. 5 leads to the *Mean Squared Error*, that is, $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i^A - \hat{y}_i)^2$.

- In OLS, let m be the number of fitted parameters. Taking $\varphi(x) = \frac{n}{n-m}x^2$ in Eq. 5 leads to the *reduced Chi-square statistic*,[10] that is, $\chi^2_{n-m} = \frac{\sum_{i=1}^{n}(y_i^A - \hat{y}_i)^2}{n-m}$.

It is worth recalling that the MSE, which measures the average of the squares of the errors, is one of the most widely used loss functions in statistics and that χ^2_{n-m} is the mechanism by which classical regression inference (t and F testing) is done since $SSE/\sigma^2 \sim \chi^2_{n-m}$, where SSE is the residual sum of squares. Note also that if the means of Y^A and Y^R are different -for example in the case of nonparametric stochastic smoothing for small samples-, we could use the indices derived in Theorem 2 to obtain similar results.

4 Conclusions

The divergence approach presented in this paper offers several significant advantages. Firstly, it provides a new framework that encompasses various economic domains, including distributional economics (inequality, poverty, richness, inequality of opportunity, and income mobility), wage discrimination, and goodness of fit. Thus, a number of indices proposed in the economic literature can be considered as particular cases of the class of divergence indices derived here. For inequality, we find, among others, the coefficient of variation, the Robin Hood index, the variance, the Kolm-Pollack family, a subfamily of consistent unit indices proposed in Zheng (2007a) and the family of indices of absolute and relative trade-offs characterized by Ebert (1988). For poverty, we have the FGT family of relative and absolute poverty indices, the absolute poverty index proposed by Zheng (2007b) and the Krtscha-type intermediate poverty index. For richness, we find the Peichl-Schaefer-Schneider convex family and the index proposed by Medeiros (2006), while for income mobility, we obtain the indices characterized by Mitra and Ok (1998). Furthermore, for wage discrimination, we obtain the convex subfamily of relative and absolute indices proposed by Jenkins (1994), and the family of relative and absolute indices proposed by del Rio et al. (2011), while as goodness-of-fit indices we find the mean squared error and the OLS reduced Chi-square statistic. It is therefore clear that this unification aids in comprehending the indices proposed in the literature to measure these distinct phenomena.

Secondly, as a more general framework, the divergence approach empowers researchers to generate new families of indices, as demonstrated in Section 3.4, where we propose three new classes of inequality of opportunity measures, one relative, one absolute and one consisting of units.

A broader setting can be introduced by relaxing some of the assumptions we have made. For example, the generalized Pigou-Dalton principle enforces convexity in the aggregation of gaps. However, certain existing classes of indices, such as the concave family of wage discrimination indices proposed by Jenkins (1994) or the convex class of richness indices derived by Peichl et al. (2010), do not meet this principle. This suggests opportunities for exploring new results by considering alternative implications of gap transfers. Additionally, as mentioned earlier, decomposability rules out Gini-type indices like the one introduced by Almås et al. (2011). Thus, a promising avenue for future research could involve characterizing economic divergence indices that are not decomposable. Finally, alternative formulations of the gapping axiom could also be subject to study. However, it is important to note that the approach we propose does not need to be flawless to serve as a valuable tool for researchers

[10] The reduced Chi-squared statistic in OLS is known as *Mean Squared Weighted Deviation* in isotopic dating, and *Unit Weight Variance* in the context of weighted least squares.

who have sought a comprehensive way to describe multiple subfields of economics within a single framework.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10888-025-09677-6.

Acknowledgements The authors acknowledge funding from MCIN/AEI/10.13039/501100011033 under grants PID2019-107539GB-I00 (Lasso de la Vega) and PID2022-137352NB-C43 (Rodríguez and Salas), and from the Basque Government under project IT1697-2 (Lasso de la Vega) and Comunidad de Madrid under project H2019/HUM-5793-OPINBI-CM (Rodríguez and Salas). The authors thank Prof. Buhong Zheng, Elena Bárcena, Kristoff Bosmans, Brice Magdalou, and the rest of participants in the tenth ECINEQ meeting for their comments. The views expressed are those of the authors not the funders.

Author Contributions All authors conceived the presented idea, developed the theoretical formalism, discussed the results, and contributed to the final version of the manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aczél, J.: Lectures on functional equations and their applications, Academic press (1966)
Aczél, J., Dhombres, J.: Functional equations in several variables, Cambridge University press (1989)
Almås, I., Cappelen, A.W., Lind, J.T., Sørensen, E., ø. and Tungodden, B.: Measuring unfair (in) equality. J. Public. Econ. **95**(7–8), 488–499 (2011)
Atkinson, A.B.: Measuring poverty and differences in family composition. Econ. **59**, 1–16 (1992)
Bossert, W.: Redistribution mechanisms based on individual characteristics. Math. Soc. Sci. **29**(1), 1–17 (1995)
Bossert, W., Fleurbaey, M.: Redistribution and compensation. Soc. Choice. Welfare. **13**(3), 343–355 (1996)
Bossert, W., Fleurbaey, M., et al.: Responsibility, talent, and compensation: A second-best analysis. Rev. Econ. Design. **4**(1), 35–55 (1999)
Chakravarty, S., Tyagarupananda, S.: 'The subgroup decomposable absolute indices of inequality', Quantitative Economics: Theory and Practice, Essays in Honor of Professor N. Bhattacharya. New Delhi: Allied Publishers Limited pp. 247–257 (1998)
Cowell, F.A.: Measures of distributional change: An axiomatic approach. The Rev. Econ. Stud. **52**(1), 135–151 (1985)
Cowell, F.A., Davidson, R., Flachaire, E.: Goodness of fit: an axiomatic approach. J. Business. Econ. Stat. **33**(1), 54–67 (2015)
Csiszár, I.: Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. The Annal. Stat. **19**(4), 2032–2066 (1991)
del Río, C., Gradín, C., Cantó, O.: The measurement of gender wage discrimination: the distributional approach revisited. The J. Econ. Inequal. **9**(1), 57–86 (2011)

Devooght, K.: To each the same and to each his own: A proposal to measure responsibility-sensitive income inequality. Econ. **75**, 280–295 (2008)
Ebert, U.: A family of aggregative compromise inequality measures. Int. Econ. Rev. **29**, 363–376 (1988)
Erdős, J.: 'A remark on the paper "On some functional equations" by S. Kurepa', Glasnik Mat.-Fiz. Astr. Drugtvo Mat. Fiz. Hrvatske **14**, 3–5 (1959)
Foster, J., Greer, J., Thorbecke, E.: A class of decomposable poverty measures. Econometric. **52**, 761–766 (1984)
Hamada, K., Takayama, N.: Censored income distributions and the measurement of poverty. Bulletin. Int. Stat. Inst. **47**, 617–630 (1977)
Jenkins, S.P.: Earnings discrimination measurement: A distributional approach. J. Economet. **61**(1), 81–102 (1994)
Kolm, S.-C.: Unequal inequalities. I. J. Econ. Theory. **12**(3), 416–442 (1976)
Kolm, S.-C.: Modern Theories of Justice. MIT Press (2002)
Krtscha, M.: A new compromise measure of inequality, in 'Models and Measurement of Welfare and Inequality', Springer, pp. 111–119 (1994)
Kuczma, M.: 'A survey of the theory of functional equations', Publikacije Elektrotehničkog fakulteta. Serija Matematika i fizika **130**, 1–64 (1964)
Magdalou, B., Nock, R.: Income distributions and decomposable divergence measures. J. Econ. Theory. **146**(6), 2440–2454 (2011)
Medeiros, M.: The rich and the poor: The construction of an affluence line from the poverty line. Soc. Indic. Res. **78**(1), 1–18 (2006)
Mitra, T., Ok, E. A.: 'The measurement of income mobility: A partial ordering approach', Economic Theory pp. 77–102 (1998)
Peichl, A., Schaefer, T., Scheicher, C.: Measuring richness and poverty: A micro data application to Europe and Germany. Rev. Income. Wealth. **56**(3), 597–619 (2010)
Pignataro, G.: Equality of opportunity: Policy and measurement paradigms. J. Econ. Surv. **26**(5), 800–834 (2012)
Prieto-Rodríguez, J., Rodríguez, J. G., Salas, R.: The measurement of wage discrimination with imperfect information: A finite mixture approach, In: 'Research on Economic Inequality. Inequality, Redistribution and Mobility', Emerald Publishing Limited, **28**, pp. 187–207 (2020)
Ramos, X., van de Gaer, D.: Approaches to inequality of opportunity: Principles, measures and evidence. J. Econ. Surv. **30**(5), 855–883 (2016)
Roemer, J.E., Trannoy, A.: Equality of opportunity: Theory and measurement. J. Econ. Lit. **54**(4), 1288–1332 (2016)
Sen, A.: Poverty: An ordinal approach to measurement. Economet. **44**, 219–231 (1976)
Shorrocks, A.F.: Inequality decomposition by population subgroups. Economet. **52**, 1369–1385 (1984)
Shorrocks, A. F.: Aggregation issues in inequality measurement, In: 'Measurement in Economics', Springer, pp. 429–451 (1988)
Zheng, B.: Poverty orderings. J. Econ. Surv. **14**(4), 427–466 (2000)
Zheng, B.: Unit-consistent decomposable inequality measures: Some extensions, Technical report, Working paper, University of Colorado, Department of Economics (2005)
Zheng, B.: Unit-consistent decomposable inequality measures. Economica. **74**, 97–111 (2007)
Zheng, B.: Unit-consistent poverty indices. Econ. Theory. **31**(1), 113–142 (2007)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.