Descriptive Units of Heterogeneity: An Axiomatic Approach to Measuring Heterogeneity

(Latest Version)

Willy Chen*

August 6, 2025†

Abstract

This paper proposes a new way of measuring system heterogeneity, highlighting the limitations of existing measures in providing cardinal interpretations and facilitating meaningful comparisons across contexts. Using an axiomatic framework, I evaluate the strengths and weaknesses of standard inequality and concentration measures and identify general properties that heterogeneity measures should satisfy. Building on these principles, I propose the Descriptive Units of Heterogeneity (DUH)—a family of measures designed specifically for contexts where the primary objective is to assess distributional heterogeneity rather than deconcentration. DUH retain the broad comparability of concentration measures while offering greater sensitivity to changes among smaller groups. Two empirical illustrations—one involving racial composition in cities and another examining firm-level revenue diversification—demonstrate the practical advantages of DUH in capturing nuanced structural changes.

Keywords: Diversity, Concentration, Inequality, Heterogeneity
JEL: B41, D30, D63, J15, L11

*Michigan State University: willyc@msu.edu

†I thank Jon Eguia, Ce Liu, Sarah Mirrow, Hanzhe Zhang, and referees for suggestions and feedback.

1 Introduction

The problem of quantifying and ranking heterogeneity across systems—particularly in the context of categorical outcomes in a population—has long posed methodological challenges. Conventional measures such as the Gini coefficient and the Herfindahl-Hirschman Index (HHI), while widely employed in the measurement of inequality and concentration, exhibit properties that may be ill-suited for capturing distributional heterogeneity (Nunes, Trappenberg, and Alda, 2020; Kvålseth, 2022). This paper introduces and axiomatizes a novel family of measures, termed Descriptive Units of Heterogeneity (DUH), which are designed to satisfy desiderata more appropriate for the analysis of heterogeneity than those underpinning traditional inequality or concentration measures. The nomenclature reflects a deliberate departure from the term "index," underscoring the cardinal nature of DUH and their conceptual independence from welfare-theoretic considerations. DUH are sensitive to distributional changes, particularly among small population subgroups, and thereby offer a refined lens through which heterogeneity can be assessed in contexts where deconcentration fails to serve as a sufficient statistic.

Let a system s be defined as an ordered tuple of cardinality $|s|$, where each element is a non-negative integer. This primitive can be extended to accommodate non-negative real-valued tuples, contingent on the specification of appropriate axioms. In applied contexts, each integer represents the size of a particular subgroup in the population. The concept of heterogeneity within a system pertains to the distribution of group sizes: greater heterogeneity corresponds to more balanced group sizes, whereas homogeneity reflects the dominance of one or a few groups. For instance, a bi-partition of a population into two equally sized groups (e.g., 50–50) is considered more heterogeneous than a highly imbalanced division (e.g., 99–1). Similarly, a tri-partition such as (45, 45, 10) exhibits greater heterogeneity than (85, 5, 10), due to the more equitable distribution of individuals across groups. While intuitive comparisons can often be made between such configurations, ambiguity arises in intermediate cases—for example, comparing (55, 35, 10) and (60, 20, 20). The challenge becomes more pronounced as the number of groups increases, raising the question of how to formally characterize and compare heterogeneity across arbitrary systems.

Measuring heterogeneity fundamentally involves a dimension-reduction problem. When examining a system comprising numerous groups, the goal is to simplify the complexity while preserving essential information. Selecting appropriate dimensions to characterize heterogeneity is crucial to defining measures capable of effectively detecting changes in the system. A sensible way to proceed is to determine first the bounds of the measure, and then the ordering of systems that are neither perfectly heterogeneous nor perfectly homogeneous. The numerical bounds of any heterogeneity measure are established by designating perfect heterogeneity as any system in which all groups are the same size, and perfect homogeneity as any system in which all but one group have size 0. Subsequent tasks include (1) balancing the relative influence of large groups and the distribution of elements across small groups, and (2) clearly defining how the number of groups in the system relates to heterogeneity. These considerations are essential since, for instance, it might seem intuitive that the system $(50, 50)$ is less heterogeneous than the system $(25, 25, 25, 25)$; however, formal criteria are necessary to justify such comparisons.

In applied contexts, system heterogeneity is most commonly assessed using either dispersion or concentration measures (James and Taeuber, 1985). Dispersion measures originate from the study of income inequality. While these measures are formally invariant under permutations of outcome labels, their interpretability often presupposes a meaningful ordinal structure—an assumption that may not hold in categorical contexts. Concentration measures emerged from industrial organization (Hirschman, 1945; Herfindahl, 1950) and information theory (Sims, 2003, 2010; Pomatto, Strack, and Tamuz, 2023). These measures are designed to capture the dominance of large groups within a system, at the expense of sensitivity to variation among smaller groups.

Dispersion measures, more commonly known as measures of inequality such as the Gini coefficient, quantify inequality with the distance between the observed population distribution and a benchmark distribution. While they permit straightforward descriptive interpretations, their applicability is constrained by the Lorenz criterion, a condition analogous to second-order stochastic dominance (Atkinson, 1970; Chakravarty, 2015; Cowell, 2000; Lambert, 2002). If the Lorenz criterion is violated when comparing two systems, sensible inequality measures can be arbitrarily determined to yield conflicting rankings of the same pair of systems (Atkinson, 1970). Newbery (1970) further demonstrates that no other additively linear measure can consistently replicate the Gini ranking across all distributions.

Within the domain where the Lorenz criterion holds, dispersion measures admit a simple interpretation: higher values correspond to greater heterogeneity. While the simple interpretation of the Gini coefficient contributes to its popularity, Schwartz and Winship (1980) point out that many empirical researchers fail to account for the Lorenz criterion when using the measure to rank income inequality (pp. 2, 8), and that such failure can lead to obscured inferences (pp. 9-13). To address this limitation, Weymark (1981) characterizes a set of generalized Gini absolute and relative inequality measures that rank systems consistently given assigned weights. Yet, the normative significance of these weights remains unspecified, limiting their interpretability and empirical tractability (Weymark, 2003; Gajdos and Weymark, 2005).

Concentration measures quantify the distribution of shares across groups, with particular sensitivity to the size of dominant groups. Unlike dispersion measures, they are not constrained by the Lorenz criterion and are thus applicable across a broader domain. Prominent examples include the Herfindahl–Hirschman Index (HHI) and Shannon's entropy (SE). While these measures are effective in detecting reductions in dominance such as when a large group shrinks, they exhibit limited responsiveness to variation among smaller groups. Moreover, they typically disregard the presence of zero-valued elements (zero-groups) in a system, may carry meaningful information about the structure of heterogeneity (Kvålseth, 2022). These features render concentration measures less suitable for contexts in which heterogeneity is not solely driven by deconcentration.

Although dispersion and concentration measures were developed for distinct purposes, many have been axiomatized using common principles relevant to the study of heterogeneity. The Gini coefficient and its generalizations have been studied by Rothschild and Stiglitz (1971, 1973), Schwartz and Winship (1980), and Weymark (1981). The HHI has been axiomatized by Chakravarty and Eichhorn (1991) and Kvålseth (2022), while Shannon's entropy has been axiomatized by Nambiar, Varma, and Saroch (1992), Suyari (2004), and Chakrabarti, Chakrabarty et al. (2005). These axiomatizations are often grounded in normative frameworks related to social welfare (Cowell, 2000; Lambert, 2002; Chakravarty, 2015), making their direct applicability to the measurement of heterogeneity limited.

I axiomatize DUH specifically to assess distributional heterogeneity. These measures yield cardinal interpretation, shift focus away from large groups, account for the presence of zero-groups, and share basic structural properties with established measures. The proposed framework models heterogeneity using two determinants: the relative size of the largest group and the evenness among the remaining groups. This approach builds on the axiomatic foundations of dispersion and concentration measures while incorporating desiderata more appropriate for assessing heterogeneity. The axioms are organized into two categories: fundamental axioms, which establish baseline properties, and characterization axioms, which uniquely characterize the DUH within the broader class of heterogeneity measures.

Fundamental axioms formalize the most common desiderata across measures of heterogeneity and induce partial orders over arbitrary systems. The axioms are group symmetry (GSYM), scale invariance (INV), and the principle of transfers (PT). GSYM and INV state that two systems with the same number of groups are equally heterogeneous if one is a permutation or rescaling of the other. Under GSYM, the exact order in each system does not influence heterogeneity. Under INV, the domain of systems can be generalized to tuples of non-negative real numbers without loss.[1] PT governs how two marginally different systems should be ordered. Most standard dispersion and concentration measures satisfy these axioms (Rothschild and Stiglitz, 1971, 1973; Schwartz and Winship, 1980; Chakravarty and Eichhorn, 1991; Kvålseth, 2022; Chakravarty, 2015; Nambiar, Varma, and Saroch, 1992; Suyari, 2004; Chakrabarti, Chakrabarty et al., 2005; Cowell, 2000; Nunes, Trappenberg, and Alda, 2020).

Characterization axioms induce total orders over arbitrary systems and uniquely characterize measures. I define independence (IND), the principle of proportional transfers (PPT), contractibility (CON), and unity (UNI). IND requires that the influence of the largest group is orthogonal to that of the evenness of minority groups; PPT refines PT with cardinal interpretation; CON requires that a measure of heterogeneity decreases with the addition of zero-groups; and UNI normalizes a measure to be between 0 and 1. Together, these axioms uniquely characterize DUH as a family of measures that focus on rendering the comparison of heterogeneity between systems descriptive—cardinally interpretable and reflective of changes to the distribution of small groups. These features make DUH the ideal measures for comparing the distributional heterogeneity between arbitrary systems.

As a direct result of the fundamental axioms, the rank correlation between DUH and other measures are high. However, they often diverge in ordering specific pairs of systems, particularly when the Lorenz criterion is not satisfied due to variation among smaller groups. Conditional on the size of the largest group, DUH display greater dispersion in values than existing measures, highlighting their enhanced sensitivity to distributional nuances that conventional measures tend to overlook.

Meaningful comparisons of heterogeneity require that the elements of each system be defined according to a consistent and coherent logic. In particular, the labeling of elements in a system must reflect a reasonable and interpretable classification scheme. To illustrate the importance of this requirement, I present an example in which alternative group labels yield conflicting conclusions about heterogeneity, thereby motivating the notion of reasonable group labels as a prerequisite for valid comparisons.

To further demonstrate the applicability and nuance of DUH, I consider two empirical illustrations. First, using the evolving racial composition of a hypothetical city, I show that DUH are well-suited for contexts where the primary concern is heterogeneity itself, rather than inequality or concentration. In this example, the size of the majority group increases over time, while the distribution among minority groups becomes more balanced. DUH capture this dual movement

[1] This can be an undesirable property in contexts where the scale of the system matters such as ecological heterogeneity. See Nunes, Trappenberg, and Alda (2020) for a more complete discussion.

by reflecting both the growing dominance of the majority and the increasing evenness among minorities. In contrast, standard measures yield conflicting signals: some increase due to minority balance, while others decrease due to majority growth.

Second, I examine changes in Apple's revenue composition from 2012 to 2023. This example illustrates DUH's sensitivity to structural shifts in the distribution of smaller categories. While the size of the dominant revenue source remains relatively stable, the evenness among the remaining sources declines over time. DUH values respond markedly to this change, registering a significant decrease in heterogeneity. In contrast, standard concentration measures exhibit only modest declines. This divergence highlights DUH's capacity to detect meaningful distributional changes in the presence of a dominant group, thereby offering a more nuanced assessment of heterogeneity than conventional measures.

The remainder of the paper proceeds as follows. Section 2 introduces the foundational primitives for analyzing system heterogeneity, reviews standard measures of dispersion and concentration, and motivates the need for a novel approach tailored to distributional heterogeneity. Section 3 sets out the axiomatic foundations and presents the proposed framework, which isolates two principal determinants of heterogeneity. Section 4 defines the DUH class of measures and contrasts their properties with those of conventional inequality and concentration measures. Section 5 discusses guidelines for empirical implementation and illustrates the applicability of DUH across a range of empirical settings. Formal proofs of the characterization results are provided in the Appendix.

2 Model

2.1 Primitives

The objective of a measure of heterogeneity is to induce a total order, in terms of heterogeneity, over arbitrary systems. Formally, define a system $s = (s_1, s_2, \ldots, s_G)$ is an ordered tuple of non-negative integers, excluding the null tuple. That is, $s \in \mathbb{Z}_+^G \backslash \vec{0}$ where $G = |s|$ denotes the cardinality of the tuple, corresponding to the number of groups in the population in applied contexts. Let $\mathcal{S}$ denote the collection of all systems. For any $s \in \mathcal{S}$, define the total population as $||s||_1 = \sum_{g=1}^{G} s_g$, and the mean group size as $\mu(s) = \frac{||s||_1}{||s||}$. These quantities serve as foundational descriptors of the system's structure. As an illustrative example, consider a tri-partition of 100 individuals into groups of sizes 55, 35, 10. This yields the system $s = (55, 35, 10)$, with $|s| = 3$, $||s||_1 = 100$, and $\mu(s) = \frac{100}{3}$.

Definition 1. *A function $\Phi : \mathcal{S} \to \mathbb{R}$ is a measure of heterogeneity if, for any two systems s and s',*

$$\Phi(s) \geq \Phi(s') \iff s \text{ is weakly more heterogeneous } (\succsim) \text{ than } s'.$$

Since heterogeneity denotes the presence of mixture and homogeneity denotes the lack thereof, systems that are maximally and minimally heterogeneous are naturally defined.

Definition 2. *A system $\bar{s}$ is maximally heterogeneous if:*

$$\bar{s} = (s_1, s_2, \ldots, s_G) = \mu(\bar{s}) \cdot (1, 1, \ldots, 1).$$

Definition 3. *A system $\underline{s}$ is minimally heterogeneous (or perfectly homogeneous) if all but one element of $\underline{s}$ is 0. In other words,* $\underline{s} = (0, 0, \cdots, 0, ||\underline{s}||_1, 0, \cdots, 0, 0)$.

Under Definitions 2 and 3, it is convenient to restrict the range of any measure of heterogeneity to a proper subset of $\mathbb{R}$ such as $[0, 1]$ or $[0, c]$, $c \in \mathbb{R}_{++}$. Such normalization facilitates the construction of a total order over systems that are neither maximally nor minimally heterogeneous, allowing for arbitrary but principled comparisons subject to reasonable constraints (Marshall, Olkin, and Arnold, 1979).

2.2 Measures of Dispersion/Inequality

Dispersion measures compare the observed distribution to a benchmark distribution. The Gini coefficient, a well-known dispersion measure, regards a system as the rank distribution of a single outcome, typically household income, and compares that distribution to the uniform distribution. The distance between the two distributions represents the amount of heterogeneity of the system.

Definition 4 (Ordered system $\hat{s}$)**.** *Let s be an arbitrary system. An ordered system $\hat{s} = (\hat{s}_1, \ldots, \hat{s}_G)$ of s is the permutation of s such that $\hat{s}_1 \geq \hat{s}_2 \geq \cdots \geq \hat{s}_G$.*

Blackorby and Donaldson (1978) shows that the Gini coefficient of s is identical to that of $\hat{s}$, which can be written as:

$$Gini(s) = 1 - \frac{1}{G^2} \cdot \left(\frac{\sum_{g=1}^{G} (2g-1)\hat{s}_g}{\mu(\hat{s})} \right).$$

The Gini coefficient evaluates inequality through the relationship: "the top $x\%$ of households earn the top $y\%$ of income." Under perfect equality—corresponding to minimal heterogeneity—this relationship satisfies $x = y$, $\forall x \in [0, 100]$. Deviations from this equality indicate the presence of inequality, i.e., heterogeneity in household income. While this interpretation is intuitive and widely used, it relies on a restrictive assumption: the Lorenz criterion. Specifically, the Gini coefficient provides a consistent ordering of inequality between two distributions if and only if their Lorenz curves do not intersect. When Lorenz curves cross, the Gini-based ordering becomes ambiguous, limiting its applicability for comparing heterogeneity across such cases (Atkinson, 1970; Marshall, Olkin, and Arnold, 1979; Schwartz and Winship, 1980; James and Taeuber, 1985). Nunes, Trappenberg, and Alda (2020) discuss several other dispersion measures in the context of ecology, biology, and medicine. These measures exhibit other limitations which render them less practical in the context of system heterogeneity.

To escape the Lorenz criterion, Weymark (1981) characterizes a class of generalized Gini absolute inequality measures:

$$Generalized.Gini(s) = \mu(\hat{s}) - \sum_{g=1}^{G} w_g \hat{s}_g,$$

with weights w_g satisfying $w_1 \leq w_2 \leq \cdots \leq w_g$, $\sum_{g=1}^{G} w_g = 1$. Unlike the standard Gini coefficient, which yields meaningful comparisons only when the Lorenz curves of two systems do not intersect, this generalized formulation induces total orders over systems with an equal number of groups, even in the presence of intersecting Lorenz curves. However, the very flexibility that enables such generality also limits the practical appeal of these measures, as the absence of a canonical weighting scheme complicates their interpretation and empirical application (Weymark, 2003; Gajdos and Weymark, 2005).

Other dispersion measures such as the Hoover index (Hoover, 1936) and the Atkinson index (Atkinson, 1970) focus on a normative approach to interpret the distance between the observed system and the uniform distribution. The Hoover index measures the average absolute distance between each household's income and the average household income, while the Atkinson index is a generalized measure that can specify the effect of changes to the distribution among low-income households via an inequality-aversion parameter.

Cowell (2000) offers a comprehensive treatment of inequality measures, detailing their normative foundations, technical properties, and empirical relevance. Lambert (2002) further explores their application in the context of taxation and social welfare, emphasizing the role of inequality aversion embedded in the underlying social welfare function (p. 6). Crucially, such aversion presupposes a natural ordering of groups, which aligns well with the normative objectives of inequality analysis. However, this assumption becomes problematic in the context of heterogeneity, where group identities may lack a meaningful or consistent ordering. Consequently, a normative framework grounded in inequality aversion may be less suited for capturing the structural features of heterogeneity.

2.3 Measures of Concentration

Concentration measures quantify the extent to which a system is dominated by its largest groups, thereby capturing the degree of dominance or aggregation within the distribution.

Definition 5 (P_g, $\hat{P}_g$, $\tilde{P}_g$)**.** *For any system s and its ordered system $\hat{s}$ such that $|s| = G$, define:*

$$P_g = \frac{s_g}{||s||_1}, \quad \hat{P}_g = \frac{\hat{s}_g}{||s||_1}, \quad \textit{and} \quad \tilde{P}_g = \frac{\hat{s}_g}{||s||_1 - \hat{s}_1}.$$

In contrast, measures of heterogeneity are concerned with the overall dispersion or spread across all groups, rather than the relative size of the largest ones (Kvålseth, 2022). Chakravarty and Eichhorn (1991) show that both the HHI and SE are derivatives of the *Hannah-Kay* class of concentration measures (Hannah and Kay, 1977), defined with the perception parameter $\alpha (> 0)$ as:

$$H_{\alpha}^{G}(s) = \begin{cases} \left[\sum_{g=1}^{G} P_g^{\alpha} \right]^{\frac{1}{\alpha - 1}} & \text{if } \alpha \neq 1 \\ \prod_{g=1}^{G} P_g^{P_g} & \text{if } \alpha = 1 \end{cases}$$

Consider a system s with G groups. HHI (Hirschman, 1945; Herfindahl, 1950) and its complement, the Gini-Simpson index (GSI) (Gini, 1912; Simpson, 1949), of system s are defined as:

$$HHI\left(s\right)=\sum_{g=1}^{G}\left(P_{g}\right)^{2}=H_{2}^{G}\left(s\right), \quad GSI\left(s\right)=1-HHI\left(s\right).$$

SE is defined as:

$$SE\left(s\right)=-\sum_{g=1}^{G}\left[P_{g}\cdot ln\left(P_{g}\right)\right]=-ln\left(H_{1}^{G}\left(s\right)\right).$$

HHI induces a total order over systems and admits a cardinal interpretation: it represents the probability that two independent draws, with replacement, from the population will belong to the same group. Despite its intuitive appeal, HHI places disproportionate weight on changes in the sizes of larger groups. This property is advantageous in contexts where concentration is of primary interest, such as in the assessment of market power among firms (Hirschman, 1945; Herfindahl, 1950; Chakravarty and Eichhorn, 1991) or the influence of political parties (Laakso and Taagepera, 1979).

For example, consider the two systems (48,47,5) and (60,29,11). The GSI assigns values of 0.5462 and 0.5438 to these systems, respectively—suggesting a nearly identical degree of deconcentration. However, the underlying group structures differ in a substantively meaningful way: the first system comprises two dominant groups and a small residual group, while the second features a single dominant group with a more gradual decline in size among the remaining groups. This illustrates a case in which reliance on GSI may obscure important structural distinctions, limiting its informativeness as a measure of heterogeneity. In contrast, the DUH values for these systems are 0.2865 and 0.3170, respectively, indicating that the second system is substantially more heterogeneous. This example underscores the importance of restricting the use of HHI and GSI to contexts concerned with (de-)concentration rather than employing them as general-purpose measures of heterogeneity (Kvålseth, 2022). The simplicity of interpretation offered by HHI and GSI can, in such cases, downplay salient differences between systems.

SE is another widely used measure of concentration, particularly in the information theory and rational inattention literature (Sims, 2010, 2003; Pomatto, Strack, and Tamuz, 2023). As a negative logarithmic transformation of $H_{1}^{G}\left(s\right)$, SE diverges from HHI in two key respects. First, SE lacks the cardinal interpretation of HHI. Second, SE responds differently to refinements in group structure, especially when existing groups are subdivided, highlighting its sensitivity to the granularity of classification.

Moreover, HHI, GSI, and SE are invariant to the concatenation of zero-groups to a system. While such invariance is intuitive in contexts involving market shares or uncertainty, it becomes problematic in contexts of assessing heterogeneity, where the presence of zero-groups can carry structural significance. Instead, GSI satisfies the so-called small-subgroups property, as shown by Chakravarty (2015). This property states that "the addition of a subgroup with population size smaller than that of the smallest of the existing subgroups, all other subgroups' population sizes held constant, will increase the value of the fractionalization index" (pp. 112–113). However, this result holds only when the added subgroups are non-empty, due to the measure's invariance to zero-groups. In contrast, DUH satisfies the small-subgroups property and responds to the addition of zero-groups by decreasing in value. This added sensitivity enhances DUH's capacity to compare systems with cardinality and to detect structural changes that remain obscured under measures exhibiting zero-group invariance.

The precise ordering of heterogeneity across systems depends critically on the nature of the comparison and the degree of similarity between the systems under consideration. Contrary to the common idiom, comparisons between fundamentally different entities—such as apples and oranges—need not be inherently meaningless. When evaluated along well-defined and shared dimensions (e.g., density, color intensity, or sugar concentration), such comparisons can yield informative insights. It is even conceivable to assess both entities relative to a property characteristic of only one—for instance, citrus quality—though such an evaluation would predictably disadvantage the non-citrus item. However, certain comparisons remain intrinsically problematic. For example, comparing x to y when asserting that an apple is x times denser than an ice cube, and an orange is y times more round than a bowling ball.

To meaningfully compare heterogeneity across systems in practice, it is essential to first establish the criteria that render said systems comparable. A key consideration is that the two systems use 'comparable' group labels, including those associated with zero population. Incorporating zero-groups in a measure provides a consistent reference frame for evaluating heterogeneity. Accordingly, when comparing systems with differing numbers of groups, the system with fewer groups should be interpreted as implicitly containing additional groups with zero population, such that the two system have the same number of groups. This approach ensures that heterogeneity is assessed over a common group structure, thereby enabling valid and interpretable comparisons.

Adapting a concentration measure to account for the inclusion of zero-groups is hardly revolutionary. Cracau and Lima (2016) show that a normalized version of HHI solves this issue by revising the formula to:

$$NHHI\,(s) = \frac{HHI\,(s) - \frac{1}{G}}{1 - \frac{1}{G}} \in [0, 1].$$

This measure improves system comparability by including the number of groups in its functional form at the expense of HHI's cardinal interpretation. [2]

3 Axioms for Measures of Heterogeneity

As discussed above, the ordering of heterogeneity across systems can, in many cases, be determined flexibly to accommodate specific properties subject to some limitations. In this section, I propose a set of axioms that reflect desiderata more appropriate for the measurement of distribu-

[2]Consider the following two systems:

$$s = (0.4, 0.4, 0.2) \text{ and } s' = (0.5, 0.3, 0.1, 0.1).$$

These two systems have the same HHI (0.36), but they have different NHHIs ($NHHI(s, G = 3) = 0.04$ and $NHHI(s', G = 4) \approx 0.15$). By observation, it may not be clear whether s and s' are equally homogeneous, but a comparison of these two NHHIs is unlikely to be convincing. Once we account for zero-groups and concatenate s to make $s'' = (0.4, 0.4, 0.2, 0)$, the NHHIs of the two systems are the same (0.15)—just like their HHIs—but the level of heterogeneity can no longer be intuitively interpreted.

tional heterogeneity. These axioms are intended to guide the construction and evaluation of heterogeneity measures that are sensitive to structural features beyond mere dispersion or concentration.

I begin by introducing the set of fundamental axioms that induce a partial order over arbitrary systems. These axioms establish conditions under which two systems should be regarded as equally heterogeneous, and how two systems should be ordered when one is derived from the other through an elementary transformation.

Axiom 1 (Group Symmetry (GSYM))**.** *A measure of heterogeneity Φ satisfies the property of Group Symmetry if for every permutation $\pi(s)$ of s, $\Phi(s) = \Phi(\pi(s))$.*

GSYM requires that the heterogeneity of a system is invariant to permutations. For example, take $s_1, s_2, s_3 \in \mathbb{R}_+$,

$$\begin{array}{ccccc} s & & s' & & s'' \\ \begin{array}{|c|}\hline s_1 \\ \hline s_2 \\ \hline s_3 \\ \hline\end{array} & \sim & \begin{array}{|c|}\hline s_2 \\ \hline s_1 \\ \hline s_3 \\ \hline\end{array} & \sim & \begin{array}{|c|}\hline s_3 \\ \hline s_2 \\ \hline s_1 \\ \hline\end{array} \end{array}$$

GSYM ensures that a measure of heterogeneity depends solely on the distribution of group sizes, not on the specific labels assigned to those groups. If two systems have the same number of groups and their group size distributions are permutations of each another, they should be assigned the same heterogeneity value. Satisfying GSYM allows for meaningful comparisons between systems defined over different sets of group labels, given equal cardinality of the two sets. This axiom abstracts away from group identity and focuses exclusively on structural features of the distribution.

Axiom 2 (Scale Invariance (INV))**.** *A measure of heterogeneity Φ satisfies the property of Scale Invariance if for any system S and a scalar $\lambda \in \mathbb{R}_{++}$, $\Phi(s) = \Phi(\lambda \cdot s)$.*

INV requires that heterogeneity depends solely on the distribution of relative group sizes, not on their absolute magnitudes. Consequently, the measure is invariant to unit conversions (e.g., from count to shares). Importantly, a measure of heterogeneity satisfying INV expands its domain from ordered tuples of non-negative integers to ordered tuples of non-negative real numbers, facilitating broader applicability and interpretability.

Axiom 3 (Principle of Transfers (PT))**.** *Let $\hat{s}$ be an arbitrary ordered system. Let e_i^G be an ordered tuple of length G such that its i^{th} element is 1 and the rest are 0. A measure of heterogeneity Φ satisfies the Principle of Transfers if $\forall i < j \leq G$,*

$$\begin{cases} \hat{s}_i - \hat{s}_j \geq 2 \\ \hat{s}_i - \hat{s}_{i+1} \geq 1 \\ \hat{s}_{j-1} - \hat{s}_j \geq 1 \end{cases} \quad \textit{together implies } \Phi(\hat{s}) < \Phi(\hat{s} - e_i + e_j).$$

PT requires that, holding the order of group sizes constant, a transfer from a larger group to a smaller group increases heterogeneity of the system. For example, take the following ordered systems $\hat{s}$, $\hat{s}'$, $\hat{s}''$ with some $\varepsilon < \min\left\{s_2 - s_3, \dfrac{s_1 - s_2}{2}\right\}$. The ranking of heterogeneity of the ordered systems is:

$$
\begin{array}{c}
\hat{s} \\
\begin{array}{|c|}
\hline
\hat{s}_1 \\
\hline
\hat{s}_2 \\
\hline
\hat{s}_3 \\
\hline
\end{array}
\end{array}
\prec
\begin{array}{c}
\hat{s}' \\
\begin{array}{|c|}
\hline
\hat{s}_1 - \varepsilon \\
\hline
\hat{s}_2 + \varepsilon \\
\hline
\hat{s}_3 \\
\hline
\end{array}
\end{array}
\prec
\begin{array}{c}
\hat{s}'' \\
\begin{array}{|c|}
\hline
\hat{s}_1 - \varepsilon \\
\hline
\hat{s}_2 \\
\hline
\hat{s}_3 + \varepsilon \\
\hline
\end{array}
\end{array}
\equiv \quad \Phi(\hat{s}) < \Phi(\hat{s}') < \Phi(\hat{s}''),
$$

for any measure satisfying PT. Here, $\hat{s}'$ is obtained from $\hat{s}$ by transferring ε from the largest group to the second largest, and $\hat{s}''$ is obtained by transferring ε from the largest group to the smallest. In both cases, PT requires that heterogeneity increases: $\hat{s} \prec \hat{s}' \prec \hat{s}''$.

This axiom was first formalized by Dalton (1920) in the context of income inequality, stating: "If there are only two income receivers, and a transfer occurs from the richer to the poorer, inequality is reduced" (p. 351). It is commonly referred to as the Pigou–Dalton transfer principle, and has long been regarded as a foundational requirement for inequality measures (Pigou, 1912; Dalton, 1920; Weymark, 1981).

Kolm (1976) shows that any heterogeneity measure Φ satisfying GSYM, INV, and PT induces a total ordering over systems that satisfy the Lorenz criterion. That is, all measures satisfying GSYM, INV, and PT agree in their ordering of systems whose Lorenz curves do not intersect. However, when the domain is extended to include systems whose Lorenz curves intersect, the axioms of GSYM, INV, and PT are no longer sufficient to induce a total order. In such cases, additional axioms, reflecting more refined normative judgments, are necessary to fully characterize the heterogeneity ordering across systems.

3.1 The Two Determinants of Heterogeneity

The Gini coefficient, HHI, and SE satisfy GSYM because their functional forms treat all groups identically. However, GSYM can be satisfied under a more nuanced interpretation: by treating each *type* of groups, rather than each group, identically. In any system, there is always a largest group and a set of remaining groups. A measure that distinguishes the largest group from the rest—while treating all groups of the same type symmetrically—still satisfies GSYM.

This perspective introduces a useful decomposition of heterogeneity into two distinct determinants: (1) the dominance of the largest group, and (2) the distribution among the remaining groups. If a heterogeneity measure is constructed such that the influence of the largest group is orthogonal to that of the rest, then changes in the overall measure can be interpreted with equivalent changes in either component while holding the other constant. This decomposition enhances interpretability and allows for more granular analysis of structural variation across systems.

For clarity and convenience in all that follows, I refer to the group with the largest population share in a system as the largest group, and all other groups as minority groups.

Under INV, normalizing a system to an ordered tuple of group shares does not affect the measured heterogeneity of the system. Under this framework, a measure Φ satisfying GSYM and INV can thus be expressed as a composite function $\Phi = \Phi(\phi, \psi)$, where $\phi\left(\hat{s}\right)$ captures the influence of the largest group and $\psi\left(\hat{s}\right)$ captures the contribution of minority groups. This decomposition supports axioms that separately characterize the structural roles of dominant and subordinate groups, allowing greater flexibility in functional form.

This approach also opens the door to alternative definitions of group types, enabling the construction of group-symmetric measures without requiring identical treatment of all groups in the functional form. The specific decomposition adopted here allows for clear and interpretable assessments of system heterogeneity. In particular, the role of the largest group in this framework is analogous to that of a numeraire good in classical economics: it serves as a reference point against which the distribution of the remaining groups is evaluated. This composition allows each system to be meaningfully compared to another system with either the same largest group size or the same distribution among minority groups, thereby facilitating more nuanced and flexible comparisons of heterogeneity.

The characterization axioms introduced below, together with the fundamental axioms, induce a total order over arbitrary systems and uniquely characterize the DUH family. Like existing axiomatizations of HHI and SE, this approach emphasizes cardinal comparability, but the type-of-group framework allows for finer distinctions by isolating the dual sources of heterogeneity.

Axiom 4 (Independence (IND))**.** *A measure of heterogeneity Φ satisfies Independence if it is a composite function $\Phi(\phi, \psi)$ where $\phi(\hat{s}) = \phi(\hat{s}_1, \hat{s}_2, \ldots \hat{s}_G)$ and $\psi(\hat{s}) = \psi(\hat{s}_2, \ldots, \hat{s}_G)$*

IND requires that the contributions of the largest group and the minority groups to overall heterogeneity be treated independently. In particular, it allows for the interpretation of any changes in heterogeneity to equivalent changes resulting from adjustments solely to either the size of the largest group or the distribution among the minority groups.

Definition 6. *Take $p \in [1, \infty)$. The function $\psi_p : \mathbb{Z}_+^{G-1} \setminus \{\vec{0}\} \to \mathbb{R}$ defined as:*

$$\psi_p(\hat{s}) = \psi(\hat{s}_2, \hat{s_3}, \ldots \hat{s}_G) = 1 - \left(\sum_{g=2}^{G} \left| \tilde{P}_g - \frac{1}{G-1} \right|^p \right)^{\frac{1}{p}},$$

is a measure of evenness in the distribution of minority groups.

To capture features of the minority groups' distribution with a function ψ_p, I define it to be a function of the L^p-norm between the observed distribution in the minority groups and the uniform distribution in the minority groups. The L^p-norms with $p \in [1, \infty)$ and the vector of the $G-1$ minority groups form a well-defined metric space. Although the L^0-norm constitutes a valid metric on $\mathbb{Z}^n$ and $\mathbb{R}^n$, a function that merely indicates whether the minority distribution deviates from uniformity lacks sufficient granularity to be informative. Since $\tilde{P}_g \in [0, 1]$ and $\sum_{g=2}^{G} \tilde{P}_g = 1$ by construction, the L^p-norm between the minority distribution and the uniform distribution is always between zero and unity. The specific affine transformation in Definition 6 on the L^p-norm makes it such that $\psi_p(\hat{s}) = 0$ corresponds to a system with a minimally heterogeneous minority distribution and $\psi_p(\hat{s}) = 1$ corresponds to a system with a maximally heterogeneous minority distribution.

In pursuit of some cardinal interpretation of changes in Φ when ψ_p is fixed, I introduce the principle of proportional transfers—a minimalist refinement of PT that builds on IND.

Axiom 5 (Principle of Proportional Transfers (PPT))**.** *Let $\hat{s}$ be an arbitrary ordered system. Let e_i^G be an ordered tuple of length G such that its i^{th} element is 1 and the rest are 0. A measure of heterogeneity Φ of $\hat{s}$ satisfies the Principle of Proportional Transfers if $c \in \mathbb{R}_+,\ \alpha \in \mathbb{R}_{++}$*

$$\frac{\hat{s}_1 - c}{\hat{s}_1 + \hat{s}_2 + \cdots + \hat{s}_G} = \left(\frac{\hat{s}_1}{\hat{s}_1 + \hat{s}_2 + \cdots + \hat{s}_G} \right)^{\alpha} \text{ and } \hat{s}_1 - c \geq \hat{s}_2 + \tilde{P}_2 \cdot c$$

$$\Rightarrow \Phi\left(\hat{s} - c \cdot e_1^G + \sum_{g=2}^{G} \tilde{P}_g c \cdot e_g^G\right) = \alpha \cdot \Phi\left(\hat{s}\right)$$

PPT specifies that, holding the order of size of groups constant, a proportional transfer from the largest group to the minority groups—such that the size of the largest group is reduced from $\hat{P}_1$ to $\left(\hat{P}_1\right)^\alpha$, for some $\alpha \in \mathbb{R}_{++}$—should increase heterogeneity by a factor of α. In other words, any measure satisfying PPT responds to reductions in the size of the largest group in a tractable and predictable manner, provided that (1) the distribution among the minority groups remains unchanged, and (2) the original largest group retains its status as the largest group after the transfer. This axiom formalizes the intuition that diminishing dominance, while preserving internal minority structure, increases heterogeneity in an interpretable way. Take the comparison of heterogeneity between s and s' such that

$$\begin{array}{c} \hat{s} \\ \begin{array}{|c|} \hline \hat{P}_1 \\ \hline \hat{P}_2 \\ \hline \hat{P}_3 \\ \hline \end{array} \end{array} \prec \begin{array}{c} \hat{s}' \\ \begin{array}{|c|} \hline \hat{P}_1^\alpha \\ \hline \hat{P}_2 + \frac{\hat{P}_2}{\hat{P}_2+\hat{P}_3}\left(\hat{P}_1 - \hat{P}_1^\alpha\right) \\ \hline \hat{P}_3 + \frac{\hat{P}_3}{\hat{P}_2+\hat{P}_3}\left(\hat{P}_1 - \hat{P}_1^\alpha\right) \\ \hline \end{array} \end{array}.$$

A measure Φ satisfying GSYM, INV, and PPT would yield:

$$\Phi\left(\hat{s}\right) < \alpha\Phi\left(\hat{s}\right) = \Phi\left(\hat{s}'\right).$$

PPT endows changes in the heterogeneity measure a cardinal interpretation. Specifically, the relationship $2\Phi\left(s\right) = \Phi\left(s'\right)$ can be interpreted as "s' is twice as heterogeneous as s." This interpretation holds because s' has the same value of heterogeneity as would result from reducing the largest group proportion in s from $\hat{P}_1$ to $\left(\hat{P}_1\right)^2$, while maintaining both the evenness of the minority group distribution and the dominance of the original largest group. Thus, PPT provides a tractable and interpretable link between proportional reductions in group dominance and proportional increases in measured heterogeneity.

Axiom 6 (Contractibility (CON))**.** *Let s be an arbitrary system. Let s' be the concatenation of s and the tuple (0) such that $s' = (s_1, s_2, \ldots, s_G, 0)$. Let $\hat{s}$ and $\hat{s}'$ denote the ordered systems of s and s'. A measure of heterogeneity Φ satisfies Contractibility if*

$$\hat{s}_2 > 0 \;\Rightarrow\; \Phi\left(\hat{s}'\right) < \Phi\left(\hat{s}\right).$$

CON requires that, upon extending a system by concatenating a system with a new element 0, the overall heterogeneity of the system must decrease. A key implication of this axiom is that any comparison of heterogeneity across systems implicitly assumes a common group structure, even if some groups are unpopulated. For example, the distribution $\left(\frac{1}{2}, \frac{1}{2}\right)$ is maximally heterogeneous among two-group systems, while $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ is maximally heterogeneous among three-group systems. By CON, comparing these two requires interpreting the former as $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ which is less heterogeneous than $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

Axiom 7 (Unity (UNI))**.** *Take an arbitrary ordered system $\hat{s}$. A measure of heterogeneity Φ satisfies Unity if*

$$\Phi\left(\hat{s}\right)=0 \iff \hat{P}_1=1$$
and
$$\Phi\left(\hat{s}\right)=1 \iff \hat{P}_1=\hat{P}_2=\cdots=\hat{P}_G=\frac{1}{G}$$

The axiom UNI serves as a normalization condition, requiring that the heterogeneity measure $\Phi\left(\hat{s}\right)$ attains the value 0 for a minimally heterogeneous system and 1 for a maximally heterogeneous system. A measure satisfying UNI thus enables meaningful baseline comparisons, allowing the heterogeneity of any given system to be evaluated relative to these two reference points.

The Gini coefficient, the HHI, and SE satisfy all of the fundamental axioms as well as some of the characterization axioms. The additional axioms used to characterize HHI and SE are listed in Table 1 and formally defined in Appendix C.

4 Characterizing The Descriptive Units of Heterogeneity

Proposition 1. *Let s be an arbitrary system with four or more groups. Consider a measure of heterogeneity $\Phi_p\left(s\right)=\Phi\left(\phi\left(s\right),\psi_p\left(s\right)\right)$ that satisfies GSYM, INV, and IND. Holding $\hat{P}_1$ constant, if Φ is strictly increasing in ψ_p, then $\Phi\left(\phi,\psi_p\right)$ satisfies PT if and only if $p>1$.*[3]

Proposition 2. *If a measure of heterogeneity $\Phi(\phi,\psi)$ satisfies GSYM, INV, IND, and PPT, then it must be $\Phi=\phi\cdot\psi_p$ where $\phi\left(\hat{s}_1,\ldots\hat{s}_G\right)=-c\cdot log_q\left(\hat{P}_1\right)$, $c,q\in\mathbb{R}_{++}$.*

Definition 7. *Let $\hat{s}=\left(\hat{s}_1,\hat{s}_2,\ldots,\hat{s}_G\right)$ be an arbitrary ordered system of s. Denote $\hat{P}_1=\frac{\hat{s}_1}{\hat{s}_1+\hat{s}_2+\cdots+\hat{s}_G}$ and $\tilde{P}_g=\frac{\hat{s}_g}{\hat{s}_2+\cdots+\hat{s}_G}$ where $g\in\{2,\ldots,G\}$ and $p>1$. The family of descriptive units of heterogeneity (DUH) of the system s with $|s|\geq 2$ is:*

$$DUH\left(s\right)=\frac{ln\left(\hat{P}_1\right)}{ln(G)}\cdot\left[\left(\sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}-1\right].$$

Theorem 1: *The descriptive units of heterogeneity constitute a uniquely determined family of measures—up to positive scalar multiples—that incorporate evenness through the function ψ_p, and satisfy the axioms of group symmetry, scale invariance, independence, the principle of transfers, the principle of proportional transfers, and contractibility.*

Proof sketch[4]

[3] This proposition suggests that we need to be careful regarding the functional form of $\phi\left(\hat{P}_1\right)$. To satisfy PT, it must be that when there is a transfer from the largest group to a minority group, the increase in $\phi\left(\hat{P}_1\right)$ must dominate the decrease in ψ_p in the case where evenness decreases. Also, if $G=3$, this proposition holds with $p\geq 1$.

[4] Full proof in Appendix B.

Step 1: I show that any index using an ordered system divided by the population size as inputs satisfies GSYM and INV.

Step 2: I show that any $\Phi(\phi, \psi_p)$ that satisfies GSYM, INV, and PPT must be a positive monotonic transformation of $\frac{1}{\hat{P}_1}$.

Step 3: I show that if an $\Phi(\phi, \psi_p)$ satisfies INV, IND, and PPT, then Φ must be multiplicatively separable. In other words, $\Phi = \phi \cdot \psi_p$.

Step 4: I show that, holding $\hat{P}_1$ constant and assuming GSYM, INV, and IND, the measure Φ, using the measure of evenness ψ_p as defined, satisfies PT if and only if $p > 1$ when $G > 3$ and $p \geq 1$ when $G = 3$.

Step 5: I show that if $\Phi(\phi, \psi_p)$ satisfies GSYM, INV, IND, and PPT, then $\phi = -c \cdot log_q \left(\hat{P}_1 \right)$, $c, q \in \mathbb{R}_{++}$.

Step 6: For the case of $p = 2$, using the Euclidean distance for ψ_p, I show that the DUH family satisfies PT, in addition to GSYM, INV, IND, and PPT, by taking the derivative of the extreme case in which $\hat{P}_1$ is close to 1 and $\psi = 1$ with respect to a transfer from the largest group to the second-largest group.

Step 7: I show that the DUH family satisfies CON and UNI, in addition to GSYM, INV, IND, PPT, and PPT.

4.1 The Role of The Evenness Parameter p

Notice that the evenness parameter p, from the function ψ_p, determines the contribution of evenness in DUH. As p increases, minority groups that are farther away from $\frac{1}{G-1}$ are weighted more. At the extreme, as $p \to \infty$,

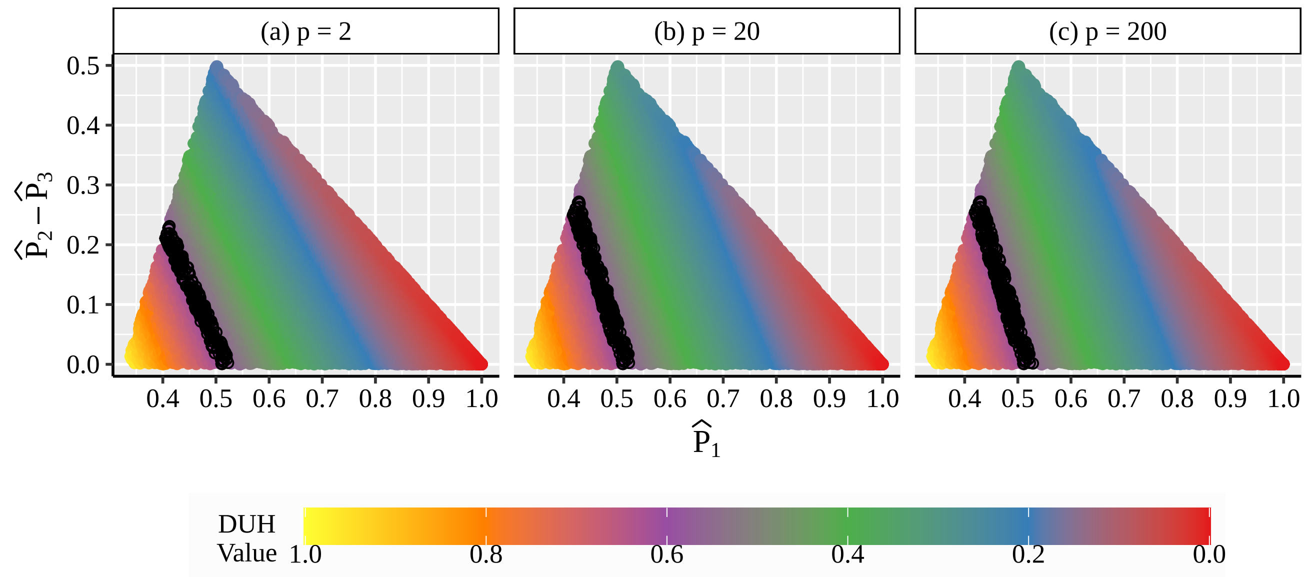
$$L^p \to L^\infty = \max_{g \in \{2, \ldots, G\}} \left\{ \left| \tilde{P}_g - \frac{1}{G-1} \right| \right\}.$$

For sufficiently large values of the evenness parameter p, the function ψ_p becomes increasingly sensitive to the most extreme deviations from evenness, such that the heterogeneity measure is effectively determined by either the largest or the smallest minority group. As illustrated in Figure 1, the variation in DUH values, conditional on a fixed value of $\hat{P}_1$, decreases with increasing p. This reflects the fact that, as the power parameter in the L^p-norm increases, ψ_p assigns greater weight to the deviation of the most extreme group share from the uniform benchmark $\frac{1}{G-1}$. The systems highlighted by the black band correspond to those with DUH values in the interval $[0.58, 0.62]$. As p increases, the slope of this band becomes steeper, indicating a reduction in the range of DUH values associated with a given $\hat{P}_1$.

4.2 Comparing DUH to Other Measures

Table 1 shows which axioms are satisfied by the Gini coefficient, DUH, HHI, and SE. Although conceptually distinct from the Gini coefficient, DUH satisfy all axioms that the Gini coefficient

Figure 1: Distribution of DUH with Different p over Systems with 3 Groups



does and more; the additional characterization axioms enable DUH to induce a total order over arbitrary systems, thereby overcoming the limitations imposed by the Lorenz criterion.

In contrast, the differences in axiomatic properties among DUH, HHI, and SE reflect the distinct normative foundations and intended applications of each measure. Specifically, HHI and SE satisfy expandability, i.e., invariance to the inclusion of zero-groups, while DUH do not. Both HHI and SE satisfy their own additional characterization axioms, endowing these measures unique cardinal interpretations of concentration.[5] Importantly, all measures considered here can be affine-transformed, for a given G, to satisfy the normalization axiom UNI. As measures of heterogeneity, DUH complement inequality and concentration measures in applied contexts by fulfilling distinct analytical objectives.

Table 1: Measures and Axioms

Type	Axiom	Gini	DUH	HHI	SE
Fundamental	Group symmetry (GSYM)	✓	✓	✓	✓
	Scale invariance (INV)	✓	✓	✓	✓
	Principle of transfers (PT)	✓	✓	✓	✓
Characterization[4]	Independence (IND)	×	✓	✓	✓
	Principle of proportional transfers (PPT)	×	✓	×	×
	Contractibility (CON)	✓	✓	×	×
	Unity (UNI)	✓	✓	×	×
	Expandability	×	×	✓	✓
	Replication principle	×	×	✓	×
	Shannon's additivity	×	×	×	✓

[5] See Appendix B for the definitions of Expandability, the Replication Principle, and Shannon's Additivity, as well as the unique characterizations these axioms provide.

When the Lorenz criterion is satisfied between two systems, the Gini coefficient, DUH, GSI, and SE induce the same order over those systems. Consequently, the rank correlations between these measures, and more generally among all Schur-convex measures, are high. Using the permutation of distinct ordered systems that are tri-partition of a population of 100, Table 2 presents the correlation and rank correlation matrices for the case where $G = 3$, an arbitrary choice for the number of groups.[6]

Table 2: Correlation Matrices of Measures over Systems with 100 Population and 3 Groups

	Correlation Matrix				Rank Correlation Matrix			
	DUH	Gini	GSI	SE	DUH	Gini	GSI	SE
DUH	1	0.957	0.832	0.874	1	0.974	0.957	0.987
Gini		1	0.947	0.953		1	0.9951	0.987
GSI			1	0.984			1	0.985
SE				1				1

However, the key contribution of DUH lies in its ability to differentiate between systems that do not satisfy the Lorenz criterion. Figure 2 illustrates this distinction by plotting the rank percentiles of the Gini coefficient, GSI, and SE against those of DUH for the same systems shown in Table 2.[7] In the axes of Figure 2, a higher rank percentile means the measure evaluates a system to be more heterogeneous. The red diagonal lines indicate systems that receive identical absolute rank percentiles under DUH and the comparison measure. Any pair of points connected by a line with a negative slope corresponds to two systems whose heterogeneity rankings are reversed between DUH and the stated measure. All measures here satisfy the axiom PT—holding the order of size of groups constant, a transfer from a larger group to a smaller group increases heterogeneity; thus, as the size of the largest group decreases, the number of distinct systems that are more heterogeneous necessarily declines. Consequently, discrepancies in rankings between DUH and the other measures are more prevalent at the lower end of the heterogeneity spectrum than at the upper end.

To elucidate the source of relative rank differences, Figure 3 presents a direct comparison between DUH and each of the Gini coefficient, GSI, and SE. For consistency across panels, the Gini coefficient, GSI, and SE are affine-transformed to satisfy UNI. All systems have three groups.

The horizontal axis in Figure 3 represents the size of the largest group in the system, denoted $\hat{P}_1$; while the vertical axis captures the difference between the second-largest and smallest group sizes, $\hat{P}_2 - \hat{P}_3$. For any fixed value of $\hat{P}_1$, the GSI exhibits limited variation along the vertical axis, indicating insensitivity to differences between minority group sizes. In contrast, the variation in SE depends on the value of $\hat{P}_1$; when $\hat{P}_1$ is sufficiently close to $\frac{1}{3}$ or 1, SE becomes nearly invariant along the y-axis. DUH, by comparison, display substantial variation along both axes and is approximately constant only when $\hat{P}_1$ and $\hat{P}_2 - \hat{P}_3$ move in opposite directions.

5 Empirical Applications of DUH

DUH lend themselves to a wide range of empirical applications. This section presents two illustrative examples that demonstrate the practical utility of DUH in applied settings. As in the previous

[6] One can verify that similar correlation patterns hold for other values of G. See Table A1 for the case with 10 groups.
[7] The analogous figure for systems with 10 groups is provided in Figure B1.

Figure 2: Rank Correlation between Measures over Systems with 100 Population and 3 Groups

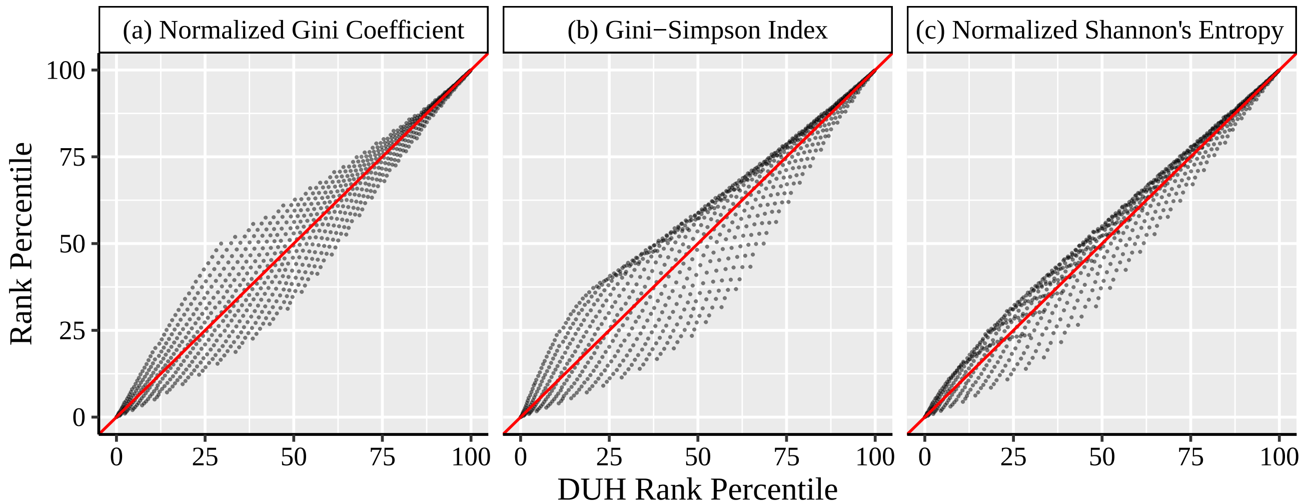
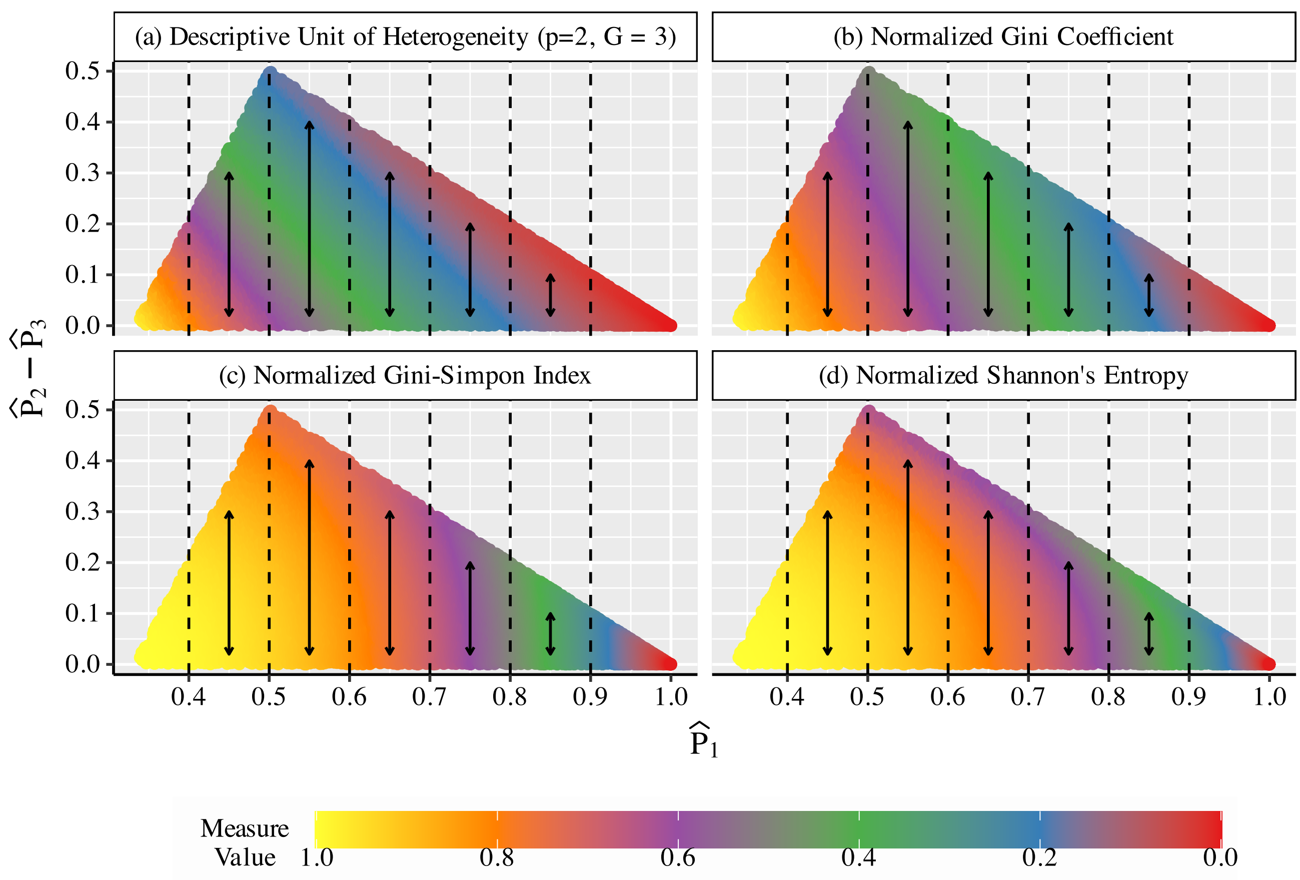


Figure 3: Comparison between DUH, Gini, GSI, and SE



section, the Gini coefficient, GSI, and SE are normalized to lie within the unit interval $[0, 1]$, wherever applicable, to facilitate direct comparison across measures.

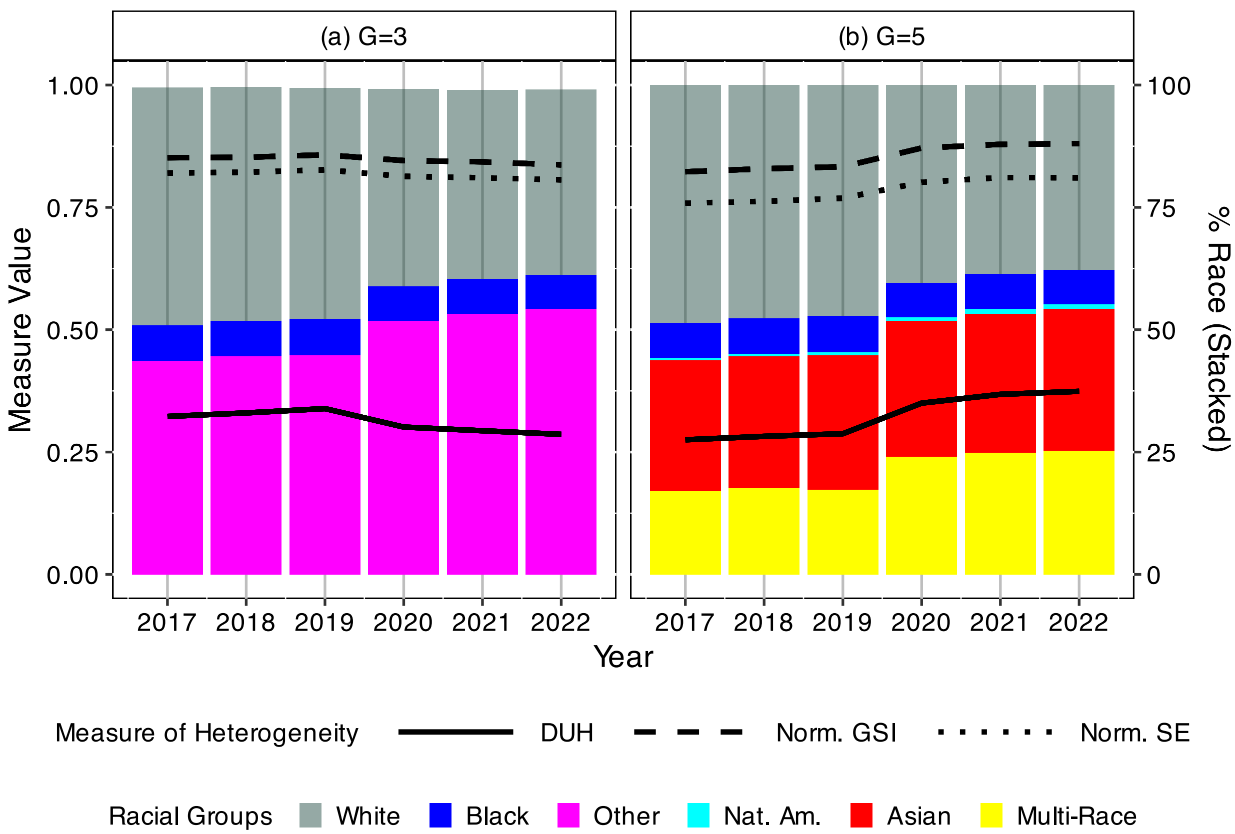
Before proceeding, it is important to emphasize the role of researcher discretion in choosing the group labels associated with a system. While systems are label-free mathematical objects, the empirical relevance and interpretability of DUH—and indeed of any heterogeneity measure—depend critically on the group labels. In all applications that follow, the DUH measure employed uses the evenness parameter $p = 2$.

5.1 Reasonable Group Labels

Satisfaction of the CON axiom implicitly requires that systems are only comparable when the labeling of elements reflect a reasonable and interpretable classification scheme. Figure 4 illustrates this principle using a practical example: the evolution of the racial composition of the San Francisco Metropolitan Statistical Area from 2017 to 2022, based on the American Community Survey (ACS) 1-year data (Ruggles et al., 2024).

In panel (a), the group labels associate with the systems are White, Black, and Other, aggregating all non-White, non-Black individuals into a single residual category. In panel (b), the Other category is disaggregated into three distinct subgroups, yielding White, Black, Asian, Native American, and Multi-Race.

Figure 4: Different Collection of Groups Can Yield Opposite Inferences



In panel (a), the three-group heterogeneity of San Francisco remains relatively stable over time, driven by a decline in the White population and a corresponding increase in the Other population. Heterogeneity begins to decline after 2019, when the Other group surpasses the White population as the largest group and continues to grow. This shift highlights the importance of the GSYM axiom, which enables researchers to interpret heterogeneity as a distributional property independent of group labels.

However, the inference changes once the group labels are redefined to capture finer distinctions among subgroups. In panel (b), where the Asian, Native American, and Multi-Race populations are considered separately, heterogeneity increases after 2019, reflecting the proportional growth of these subgroups while shrinking the White population, holding the order of group sizes constant. Such distributional changes are precisely what the PPT axiom is designed to capture.

As this example illustrates, there is no universally correct specification of group labels. The determination of group labels is a framing decision left to the discretion of the researcher. Just as the use of the Gini coefficient presupposes the Lorenz criterion, the application of any heterogeneity measure requires a justified and contextually appropriate choice of group labels. In the case shown in Figure 4, a simple refinement of a residual category alters the inference, underscoring the importance of careful consideration in defining reasonable group labels .

With these considerations in place, we now turn to examples that illustrate when and why DUH are useful. The following applications show how the DUH with $p = 2$ captures distributional features that other measures may overlook, particularly in settings where proportional differences matter.

5.2 Using DUH for Racial Heterogeneity

To demonstrate a nuanced point of DUH, let us consider the specific progression of a hypothetical city composed of three racial groups: White, Black, and Other. Table 3 reports the population shares and corresponding heterogeneity measures across five decades. Over time, the White population steadily increases, while the distribution between the two minority groups becomes more balanced.

Table 3: The Progression of Racial Composition and Racial Heterogeneity of a Hypothetical City

Share (%)	Decade				
	1	2	3	4	5
White	55%	60%	65%	66%	69%
Black	42%	34%	26%	22%	17%
Other	3%	6%	9%	11%	14%
DUH[8]	0.21	0.23	0.26	0.28	0.31
Gini[9]	0.48	0.46	0.45	0.45	0.44
Norm. GSI	0.78	0.78	0.75	0.75	0.71
Norm. SE	0.73	0.77	0.77	0.78	0.76

Under this configuration, the DUH with $p = 2$ increases monotonically, reflecting the growing evenness between the minority groups. In contrast, GSI decreases due to the rising dominance of

[8]All measure values are rounded to second digit after the decimal.

[9]Note that all systems here have intersecting Lorenz criterion so the Gini coefficient cannot be used for inference.

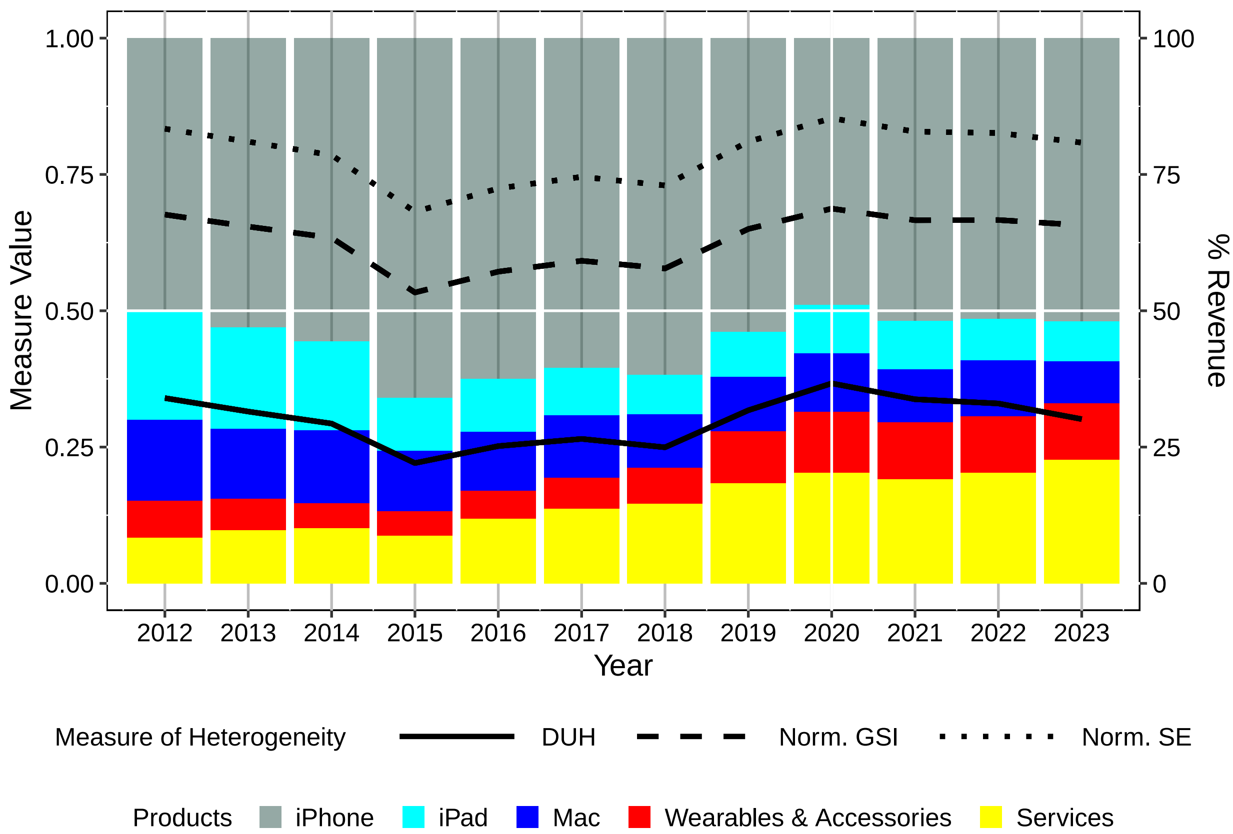
the White population. SE initially increases but declines in the final period. These divergent trends underscore the distinct properties of each measure.

The key limitation of GSI and SE is that their sensitivity to elementary transfers is highly positively correlated with group size. In this example, DUH mitigates that bias by capturing changes in the distribution among minority groups, even when those groups are relatively small. This highlights DUH's utility in contexts where minority group dynamics are of substantive interest.

5.3 Using DUH for Revenue Heterogeneity

By satisfying the axiom INV, DUH can also be applied to systems of non-negative real numbers such as a firm's revenue portfolio. Consider Apple Inc., whose revenue is generated from five major product categories: iPhone, iPad, Mac, Wearables & Accessories, and Services. Figure 5 compares DUH with GSI and SE in this context, where concentration measures are commonly used (Apple and Statista, 2024; Hirschman, 1945; Herfindahl, 1950).

Figure 5: Revenue Heterogeneity using DUH, GSI, and SE



While the three measures generally move in tandem, a notable divergence occurs between 2020 and 2023. During this period, the revenue shares of Services and Wearables & Accessories increased without a corresponding decline in the share of iPhone revenue. This shift reduced the evenness among the minority categories. DUH captures this change through a continuous and substantial decline, whereas the decrease in GSI is comparatively limited. This example highlights

DUH's sensitivity to internal distributional changes, even when the dominant category remains stable.

6 Conclusion

While measures of inequality and concentration are often employed to assess heterogeneity, their underlying value functions may introduce features that are not well-suited for capturing heterogeneity as a distributional property. To address this limitation, I characterize a new family of measures—Descriptive Units of Heterogeneity (DUH)—that evaluate heterogeneity differently from the standard repurposed measures.

Building on the axiomatic foundations of classical measures such as the Gini coefficient, the Herfindahl–Hirschman Index, and Shannon's entropy, DUH are uniquely characterized by their sensitivity to two components: the size of the largest group and the evenness among minority groups. This structure departs from symmetric algebraic treatments of all group shares and allows for a more flexible and interpretable representation of heterogeneity. DUH support cardinal interpretation of changes, including proportional statements such as "system A is x times more heterogeneous than system B," through the principle of proportional transfers.

An essential part of this explication is the importance of researcher discretion in defining the group labels associated with each system. While the overall model may appear elaborate, it provides a transparent and unambiguous framework for comparing heterogeneity across systems. As demonstrated in Section 5.2, the family DUH are not intended to replace existing measures but to complement them in contexts where sensitivity to minority group dynamics and zero-group non-invariance are analytically relevant. In settings where the focus is on dominant groups alone, such sensitivity may be undesirable. Nonetheless, DUH offer a tractable and axiomatic approach for researchers seeking to understand the evolution of heterogeneity in diverse empirical environments.

Bibliography

Apple and Statista. 2024. "Share of Apple's revenue by product category from the 1st quarter of 2012 to the 1st quarter of 2024." Graph. URL https://www.statista.com/statistics/382260/segments-share-revenue-of-apple/.

Atkinson, Anthony B. 1970. "On the measurement of inequality." *Journal of Economic Theory* 2 (3):244–263.

Blackorby, Charles and David Donaldson. 1978. "Measures of relative equality and their meaning in terms of social welfare." *Journal of Economic Theory* 18 (1):59–80. URL https://www.sciencedirect.com/science/article/pii/002205317890042X.

Chakrabarti, CG, Indranil Chakrabarty et al. 2005. "Shannon entropy: Axiomatic characterization and application." *International Journal of Mathematics and Mathematical Sciences* 2005:2847–2854.

Chakravarty, Satya R. 2015. *Inequality, Polarization and Conflict: An Analytical Study*. New Delhi: Springer India.

Chakravarty, Satya R and Wolfgang Eichhorn. 1991. "An axiomatic characterization of a generalized index of concentration." *Journal of Productivity Analysis* 2:103–112.

Cowell, Frank A. 2000. "Measurement of Inequality." In *Handbook of Income Distribution*, vol. 1, edited by A. B. Atkinson and F. Bourguignon. Elsevier, 87–166.

Cracau, Daniel and João E. D. Lima. 2016. "On the Normalized Herfindahl-Hirschman Index: A Technical Note." *International Journal on Food System Dynamics* 7 (4):382–386.

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

Gajdos, Thibault and John A Weymark. 2005. "Multidimensional generalized Gini indices." *Economic Theory* 26:471–496.

Gini, Corrado. 1912. "Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche." *Studi Economico-Giuridici della R. Università de Cagliari* 3:3–159.

Hannah, Leslie and John Anderson Kay. 1977. *Concentration in modern industry: Theory, measurement and the UK experience*. The MacMillan Press, Ltd., London.

Herfindahl, Orris C. 1950. *Concentration in the U.S. Steel Industry*. New York: Columbia University Press.

Hirschman, Albert O. 1945. *National Power and the Structure of Foreign Trade*. Publications of the Bureau of Business and Economic Research. Berkeley: University of California Press.

Hoover, Edgar M. 1936. "The Measurement of Industrial Localization." *The Review of Economics and Statistics* 18 (4):162–171. URL http://www.jstor.org/stable/1927875.

James, David R and Karl E Taeuber. 1985. “Measures of segregation.” *Sociological Methodology* 15:1–32.

Kolm, Serge-Christophe. 1976. “Unequal inequalities. II.” *Journal of Economic Theory* 13 (1):82–111.

Kvålseth, Tarald O. 2022. “Measurement of market (industry) concentration based on value validity.” *PLOS ONE* 17 (7):e0264613.

Laakso, Markku and Rein Taagepera. 1979. ““Effective” number of parties: A measure with application to West Europe.” *Comparative Political Studies* 12 (1):3–27.

Lambert, Peter. 2002. *The Distribution and Redistribution of Income: Third Edition*. Manchester: Manchester University Press.

Marshall, Albert W, Ingram Olkin, and Barry C Arnold. 1979. *Inequalities: Theory of majorization and its applications*. Springer, New York.

Nambiar, KK, Pramod K Varma, and Vandana Saroch. 1992. “An axiomatic definition of Sannon’s entropy.” *Applied Mathematics Letters* 5 (4):45–46.

Newbery, David. 1970. “A theorem on the measurement of inequality.” *Journal of Economic Theory* 2 (3):264–266.

Nunes, Abraham, Thomas Trappenberg, and Martin Alda. 2020. “The definition and measurement of heterogeneity.” *Translational Psychiatry* 10 (1):299.

Pigou, A. 1912. *Wealth and Welfare*. London: Macmillan.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2023. “The cost of information: The case of constant marginal costs.” *American Economic Review* 113 (5):1360–1393.

Rothschild, Michael and Joseph E Stiglitz. 1971. “Increasing risk II: Its economic consequences.” *Journal of Economic Theory* 3 (1):66–84.

———. 1973. “Some further results on the measurement of inequality.” *Journal of Economic Theory* 6 (2):188–204.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rodgers, and Megan Schouweiler. 2024. “IPUMS USA: Version 15.0.” dataset. URL https://doi.org/10.18128/D010.V15.0.

Schwartz, Joseph and Christopher Winship. 1980. “The Welfare Approach to Measuring Inequality.” *Sociological Methodology* 11:1–36. URL http://www.jstor.org/stable/270858.

Simpson, Edward H. 1949. “Measurement of Diversity.” *Nature* 163:688.

Sims, Christopher. 2010. “Rational Inattention and Monetary Economics.” In *Handbook of Monetary Economics*, vol. 3, edited by Benjamin M. Friedman and Michael Woodford. Amsterdam: Elsevier, 155–181.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of monetary Economics* 50 (3):665–690.

Suyari, Hiroki. 2004. "Generalization of Shannon-Khinchin axioms to nonextensive systems and the uniqueness theorem for the nonextensive entropy." *IEEE Transactions on Information Theory* 50 (8):1783–1787.

Weymark, John A. 1981. "Generalized gini inequality indices." *Mathematical Social Sciences* 1 (4):409–430. URL https://www.sciencedirect.com/science/article/pii/0165489681900184.

———. 2003. "Generalized Gini Indices of Equality of Opportunity." *The Journal of Economic Inequality* 1 (1):5–24.

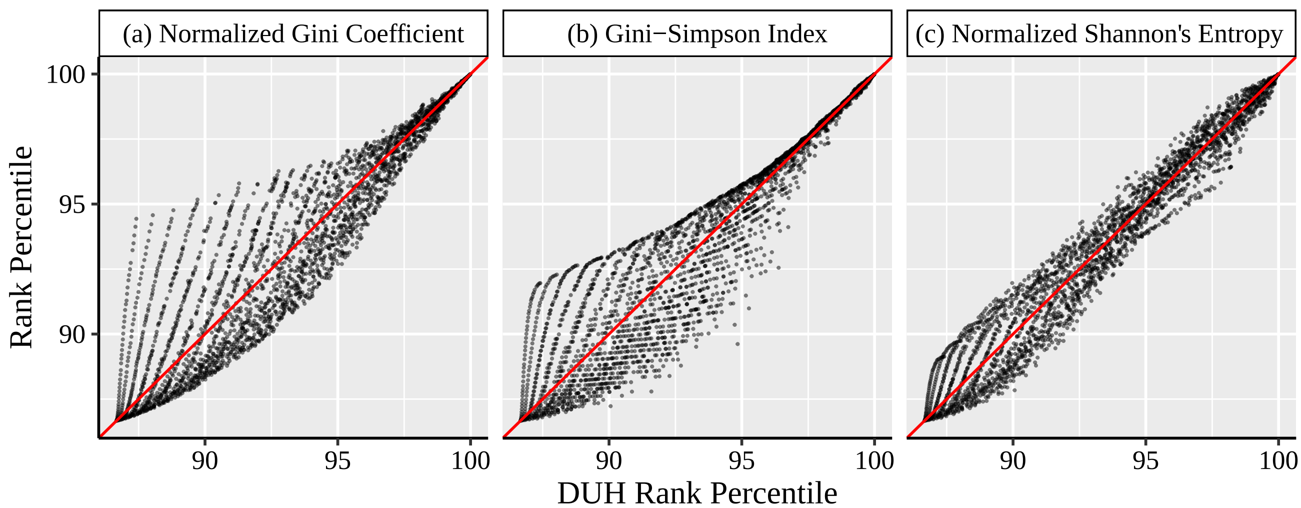
Appendices

Appendix A Additional Graphs and Tables

Table A1: Correlation Matrices of Measures over Systems with 100 Population and 3 Groups

	Correlation Matrix				Rank Correlation Matrix			
	DUH	Gini	GSI	SE	DUH	Gini	GSI	SE
DUH	1	0.910	0.760	0.888	1	0.933	0.940	0.979
Gini		1	0.559	0.688		1	0.992	0.957
GSI			1	0.947			1	0.964
SE				1				1

Figure B1: Rank Correlation of Measures over Systems with 100 Population and 3 Groups



Appendix B Proofs

Lemma 1: Any measure $\Phi(s)$ of heterogeneity satisfies GSYM if $\Phi(s) = \Phi(\hat{s})$.

Proof of Lemma 1:

The proof is trivial, given that the groups are ordered by size and not the label of the groups.

Lemma 2: If a measure of heterogeneity $\Phi(\phi, \psi)$ of satisfies GSYM, INV, and PPT, it is monotonically decreasing in $\hat{P}_1$, and therefore a positive monotonic transformation of $\frac{1}{\hat{P}_1}$.

Proof of Lemma 2:

Take any two ordered systems of G groups $\hat{s} = (\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_G)$ and $\hat{s}' = (\hat{s}'_1, \hat{s}'_2, \ldots, \hat{s}'_G)$ such that $\Phi(\hat{s}) > \Phi(\hat{s}')$ and that $(\hat{s}_2, \ldots, \hat{s}_G) = \lambda \cdot (\hat{s}'_2, \ldots, \hat{s}'_G)$, $\lambda \in \mathbb{R}_{++}$. By scale invariance:

$$\Phi(\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_G) > \Phi(\hat{s}'_1, \hat{s}'_2, \ldots, \hat{s}'_G) = \Phi\left(\hat{s}'_1 \cdot \frac{\hat{s}_S}{\hat{s}'_S}, \hat{s}'_2 \cdot \frac{\hat{s}_S}{\hat{s}'_S}, \ldots, \hat{s}'_G \cdot \frac{||s||_1}{||s'||_1}\right).$$

By the principle of proportional transfers, since $\hat{s}_1 + \hat{s}_2 + \cdots + \hat{s}_G = \hat{s}'_1 \frac{||s||_1}{||s'||_1} + \hat{s}'_2 \frac{||s||_1}{||s'||_1} + \cdots + \hat{s}'_G \frac{||s||_1}{||s'||_1}$,

$$\Phi(\hat{s}) > \Psi(\hat{s}') \iff \hat{s}_1 < \hat{s}'_1 \cdot \frac{||s||_1}{||s'||_1} \iff \hat{P}_1 < \hat{P}'_1.$$

By GSYM, this relationship is generalized to any two systems that are permutations of S and S', respectively. $\square$

Lemma 3: If a measure of heterogeneity $\Phi(\phi, \psi)$ satisfies GSYM, INV, PPT, and IND, then Φ must be multiplicatively separable, i.e., $\Phi = \phi \cdot \psi$.

Proof of Lemma 3:

Notice first that satisfying the independence axiom trivially implies that ϕ and ψ must be separable. Take any ordered system $\hat{s} = (\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_G)$. By the principle of proportional transfers, it must be that $\forall \alpha \in \left[1, \frac{\hat{s}_1 + \tilde{P}_2 \cdot \hat{s}_1}{n - 2 + \tilde{P}_2 \cdot \hat{s}_1}\right]$,

$$\begin{aligned}
\alpha \cdot \Phi\left(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_G\right) &= \Phi\left(\hat{P}_1^{\alpha}, \hat{P}_2 + \tilde{P}_2\left(\hat{P}_1 - \hat{P}_1^{\alpha}\right), \ldots, \hat{P}_G + \tilde{P}_G\left(\hat{P}_1 - \hat{P}_1^{\alpha}\right)\right) \\
\iff \underbrace{\alpha\Phi\left(\frac{\hat{P}_1}{\hat{P}_2 + \cdots + \hat{P}_G}, \tilde{P}_2, \ldots, \tilde{P}_G\right)}_{=\alpha\Phi(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_G) \text{ by INV}} &= \Phi\left(\hat{P}_1^{\alpha}, \tilde{P}_2 \cdot \lambda, \ldots, \tilde{P}_G \cdot \lambda\right) \text{ for some } \lambda \in \mathbb{R}_{++} \\
\iff \alpha\Phi\left(\left(\hat{P}_1\right), \psi\left(\tilde{P}_2, \ldots, \tilde{P}_G\right)\right) &= \Phi\left(\left(\hat{P}_1^{\alpha}\right), \psi\left(\tilde{P}_2, \ldots, \tilde{P}_G\right)\right) \\
\iff \alpha &= \frac{\Phi\left(\phi\left(\hat{P}_1^{\alpha}\right), \psi\left(\tilde{P}_2, \ldots, \tilde{P}_G\right)\right)}{\Phi\left(\phi\left(\hat{P}_1\right), \psi\left(\tilde{P}_2, \ldots, \tilde{P}_G\right)\right)}
\end{aligned}$$

For this equality to hold for all $\alpha \in \left[1, \frac{\hat{s}_1 + \tilde{P}_2 \cdot \hat{s}_1}{n-2+\tilde{P}_2 \cdot \hat{s}_1}\right]$, it must be that $\alpha = \frac{\phi\left(\hat{P}_1^{\alpha}\right) \cdot \psi\left(\tilde{P}_2, \ldots, \tilde{P}_G\right)}{\phi\left(\hat{P}_1^{\alpha}\right) \cdot \psi\left(\tilde{P}_2, \ldots, \tilde{P}_G\right)}$. Otherwise, the relationship can only hold for a specific α that is a function of the tuple $\left(\tilde{P}_2, \ldots, \tilde{P}_G\right)$.

As such, the measure Φ must be multiplicatively separable such that $\Phi = \phi \cdot \psi$. By GSYM, the result generalizes the domain from ordered systems to all systems.

□

Definition 6. *Take $p \in [1, \infty)$. The function $\psi_p : \mathbb{Z}_+^{G-1} \setminus \{\vec{0}\} \to \mathbb{R}$ defined as:*

$$\psi_p(\hat{s}) = \psi(\hat{s}_2, \hat{s}_3, \ldots \hat{s}_G) = 1 - \left(\sum_{g=2}^{G} \left|\tilde{P}_g - \frac{1}{G-1}\right|^p\right)^{\frac{1}{p}},$$

is a measure of evenness in the distribution of minority groups.

Proposition 1. *Let s be an arbitrary system with four or more groups. Consider a measure of heterogeneity $\Phi_p(s) = \Phi(\phi(s), \psi_p(s))$ that satisfies GSYM, INV, and IND. Holding $\hat{P}_1$ constant, if Φ is strictly increasing in ψ_p, then $\Phi(\phi, \psi_p)$ satisfies PT if and only if $p > 1$.*[10]

Proof of Proposition 1:

Consider two ordered systems $\hat{s} = (\hat{s}_1, \ldots, \hat{s}_g, \hat{s}_{g+1}, \ldots, \hat{s}_G)$ and $\hat{s}' = (\hat{s}_1, \ldots, \hat{s}_g - c, \hat{s}_{g+1} + c, \ldots, \hat{s}_G)$ where $c < \frac{\hat{s}_g - \hat{s}_{g+1}}{2}$. Define $\tilde{c} = \frac{c}{\hat{s}_2 + \cdots + \hat{s}_G}$. I want to show that $\psi(\hat{s}) < \psi(\hat{s}')$, meaning $\Phi(\hat{s}) < \Phi(\hat{s}')$ and that Φ satisfies *PT*.

Given $\hat{s}$ and $\hat{s}'$, we have

$$\psi_p(\hat{s}) = 1 - \left(\left|\tilde{P}_2 - \frac{1}{G-1}\right|^p + \cdots + \left|\tilde{P}_g - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} - \frac{1}{G-1}\right|^p + \cdots + \left|\tilde{P}_G - \frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}$$

$$\psi_p(\hat{s}') = 1 - \left(\left|\tilde{P}_2 - \frac{1}{G-1}\right|^p + \cdots + \left|\tilde{P}_g - \tilde{c} - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} + \tilde{c} - \frac{1}{G-1}\right|^p + \cdots + \left|\tilde{P}_G - \frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}.$$

Observe that

$$\begin{aligned}
&\psi_p(\hat{s}) < \psi_p(\hat{s}') \\
\iff &\left(\tilde{P}_g - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} - \frac{1}{G-1}\right|^p + C\right)^{\frac{1}{p}} > \left(\tilde{P}_g - \tilde{c} - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} + \tilde{c} - \frac{1}{G-1}\right|^p + C\right)^{\frac{1}{p}} \\
\iff &\left|\tilde{P}_g - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} - \frac{1}{G-1}\right|^p > \left|\tilde{P}_g - \tilde{c} - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} + \tilde{c} - \frac{1}{G-1}\right|^p.
\end{aligned}$$

Case 1: $\frac{1}{G-1} < \tilde{P}_{g+1} < \tilde{P}_g$, then

$$\left|\tilde{P}_g - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} - \frac{1}{G-1}\right|^p > \left|\tilde{P}_g - \tilde{c} - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} + \tilde{c} - \frac{1}{G-1}\right|^p$$

[10]This proposition suggests that we need to be careful regarding the functional form of $\phi\left(\hat{P}_1\right)$. To satisfy PT, it must be that when there is a transfer from the largest group to a minority group, the increase in $\phi\left(\hat{P}_1\right)$ must dominate the decrease in ψ_p in the case where evenness decreases. Also, if $G = 3$, this proposition holds with $p \geq 1$.

$$\iff \left(\tilde{P}_g - \frac{1}{G-1}\right)^p + \left(\tilde{P}_{g+1} - \frac{1}{G-1}\right)^p > \left(\tilde{P}_g - \tilde{c} - \frac{1}{G-1}\right)^p + \left(\tilde{P}_{g+1} + \tilde{c} - \frac{1}{G-1}\right)^p$$
$$\iff \frac{\left(\tilde{P}_g - \frac{1}{G-1}\right)^p + \left(\tilde{P}_{g+1} - \frac{1}{G-1}\right)^p}{2} > \frac{\left(\tilde{P}_g - \tilde{c} - \frac{1}{G-1}\right)^p + \left(\tilde{P}_{g+1} + \tilde{c} - \frac{1}{G-1}\right)^p}{2}.$$
$\iff p > 1$ (making the function x^p convex).

Case 2: $\tilde{P}_{g+1} < \frac{1}{G-1} < \tilde{P}_g$, then

$$\left|\tilde{P}_g - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} - \frac{1}{G-1}\right|^p > \left|\tilde{P}_g - c - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} + c - \frac{1}{G-1}\right|^p$$
$$\iff \left(\tilde{P}_g - \frac{1}{G-1}\right)^p + \left(\frac{1}{G-1} - \tilde{P}_{g+1}\right)^p > \left(\tilde{P}_g - c - \frac{1}{G-1}\right)^p + \left(\frac{1}{G-1} - \tilde{P}_{g+1} - c\right)^p$$
$$\iff \underbrace{\left(\tilde{P}_g - \frac{1}{G-1}\right)^p - \left(\tilde{P}_g - c - \frac{1}{G-1}\right)^p}_{>0} + \underbrace{\left(\frac{1}{G-1} - \tilde{P}_{g+1}\right)^p - \left(\frac{1}{G-1} - \tilde{P}_{g+1} - c\right)^p}_{>0} > 0.$$

Case 3: $\tilde{P}_{g+1} < \tilde{P}_g < \frac{1}{G-1}$, then

$$\left|\tilde{P}_g - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} - \frac{1}{G-1}\right|^p > \left|\tilde{P}_g - c - \frac{1}{G-1}\right|^p + \left|\tilde{P}_{g+1} + c - \frac{1}{G-1}\right|^p$$
$$\iff \left(\frac{1}{G-1} - \tilde{P}_g\right)^p + \left(\frac{1}{G-1} - \tilde{P}_{g+1}\right)^p > \left(\frac{1}{G-1} - \tilde{P}_g + c\right)^p + \left(\frac{1}{G-1} - \tilde{P}_{g+1} - c\right)^p$$
$$\iff \frac{\left(\frac{1}{G-1} - \tilde{P}_g\right)^p + \left(\frac{1}{G-1} - \tilde{P}_{g+1}\right)^p}{2} > \frac{\left(\frac{1}{G-1} - \tilde{P}_g + c\right)^p + \left(\frac{1}{G-1} - \tilde{P}_{g+1} - c\right)^p}{2}.$$
$\iff p > 1$ (making the function x^p convex).

Lastly, if $G = 3$, then only case 2 remains, meaning that the RHS of the statement can be expanded to $p \geq 1$. □

Proposition 2. *If a measure of heterogeneity $\Phi(\phi, \psi)$ satisfies GSYM, INV, IND, and PPT, then it must be $\Phi = \phi \cdot \psi_p$ where $\phi\left(\hat{s}_1, \ldots \hat{s}_G\right) = -c \cdot log_q\left(\hat{P}_1\right)$, $c, q \in \mathbb{R}_{++}$.*

Proof of Proposition 2:

From the previous two lemmas, we know that $\phi\left(\hat{P}_1\right)$ must be a positive monotonic transformation of $\frac{1}{\hat{P}_1}$ and that for α such that $\hat{P}_1^{\alpha} > \hat{P}_2 + \tilde{P}_2\left(\hat{P}_1 - \hat{P}_1^{\alpha}\right)$, we must have

$$\alpha = \frac{\phi\left(\hat{P}_1^{\alpha}\right)}{\phi\left(\hat{P}_1\right)}.$$

Notice that the only positive monotonic transformation that would satisfy this is $log_q\left(\frac{1}{\hat{P}_1}\right)$, up to a positive scalar multiplication. Further, notice that any $log_q\left(\frac{1}{\hat{P}_1}\right)$ can be rewritten as $\frac{ln\left(\frac{1}{\hat{P}_1}\right)}{ln(q)}$, so it is equivalent to write $c \cdot ln\left(\frac{1}{\hat{P}_1}\right)$. As such, $\phi\left(\hat{P}_1\right) = c \cdot ln\left(\frac{1}{\hat{P}_1}\right)$, $c \in \mathbb{R}_{++}$ is the unique function, up to positive scalar multiplication, of majority proportions that can lead to $\Phi(\phi, \psi)$ satisfying group symmetry, scale invariance, independence, and the principle of proportional transfers. $\square$

Theorem 1: The descriptive units of heterogeneity, Φ defined as:

$$\begin{aligned}\Phi_p(\hat{s}_1, \ldots, \hat{s}_G) &= -\frac{ln\left(\hat{P}_1\right)}{ln(G)}\left[1 - \left(\sum_{g=2}^{G}\left(\tilde{P}_g - \frac{1}{G-1}\right)^p\right)^{\frac{1}{p}}\right] \\ &= \frac{ln\left(\hat{P}_1\right)}{ln(G)}\left[\left(\sum_{g=2}^{G}\left(\tilde{P}_g - \frac{1}{G-1}\right)^p\right)^{\frac{1}{p}} - 1\right],\end{aligned}$$

where $\hat{P}_1 = \dfrac{\hat{s}_1}{\hat{s}_1 + \cdots + \hat{s}_G}$, $\tilde{P}_g = \dfrac{\hat{s}_g}{\hat{s}_2 + \cdots + \hat{s}_g}$, $p \in (1, \infty)$, and $\hat{s}_1 \geq \hat{s}_2 \geq \cdots \geq \hat{s}_G$,

constitute a uniquely determined family of measures—up to positive scalar multiples—that incorporate evenness through the function ψ_p, and satisfy the axioms of group symmetry, scale invariance, independence, the principle of transfers, the principle of proportional transfers, and contractibility.

Proof of Theorem 1:

Propositions 1 and 2 combined implies that DUH satisfy SYM, INV, PPT, and IND, but not necessarily PT. I have only shown that DUH satisfy PT when either ϕ or ψ_p is held constant, which means that I need to be careful with the functional form of $\phi\left(\hat{P}_1\right)$. To satisfy *PT* overall, it must be that when there is a transfer from the majority group to the minority group, the increase in Φ through $\phi\left(\hat{P}_1\right)$ dominates the decrease in Φ through ψ_p in the case in which evenness decreases as a result of the transfer.

To show that Φ satisfies PT, we only need to look at the extreme case in which $\hat{P}_1$ is close to 1 and $\psi_p = 1$. In this case, a simple transfer from $\hat{s}_1$ to $\hat{s}_2$ will decrease ψ_p the most. For simplicity, we will consider the case in which $p = 2$, so that Φ is simply

$$\Phi\left(\hat{s}_1, \ldots, \hat{s}_G\right) = \frac{ln\left(\hat{P}_1\right)}{ln(G)}\left[\sqrt{\sum_{g=2}^{G}\left(\tilde{P}_g - \frac{1}{G-1}\right)^2} - 1\right].$$

For ease of notation, denote $\hat{n}_s = ||\hat{s}||_1$ and $\tilde{n}_s = \hat{s}_2 + \cdots + \hat{s}_G = ||\hat{s}||_1 - \hat{s}_1$. A transfer of x from $\hat{s}_1$ to $\hat{s}_2$ when $\psi_2 = 1$ can be written as:

$$\Phi_2 = \frac{ln\left(\frac{\hat{s}_1 - x}{\hat{n}_s}\right)}{ln(G)}\left[\sqrt{\left(\frac{\frac{\tilde{n}_s}{G-1} + x}{\tilde{n}_s + x} - \frac{1}{G-1}\right)^2 + (G-2)\left(\frac{\frac{\tilde{n}_s}{G-1}}{\tilde{n}_s + x} - \frac{1}{G-1}\right)^2} - 1\right].$$

Taking the derivative of this expression with respect to x, we have $\forall x \in \left[0, \frac{(G-1)\hat{s}_1 - \tilde{n}_s}{G}\right]$:

$$\frac{d}{dx} ln(G)\Phi_2(\hat{s}_1 - x, \hat{s}_2 + x, \hat{s}_3, \ldots, \hat{s}_G)$$

$$= \frac{\sqrt{\frac{G-2}{G-1}}\left[\tilde{n}_s x(x-\hat{s}_1)ln\left(\frac{\hat{s}_1-x}{\hat{n}_s}\right)+x^2(\tilde{n}_s+x)\right]}{(x-\hat{s}_1)(x+\tilde{n}_s)^2}+\frac{1}{\hat{s}_1-x}$$

$$= \frac{1}{\hat{s}_1-x}\left[1-\underbrace{\frac{\sqrt{\frac{G-2}{G-1}}\left[\tilde{n}_s x(x-\hat{s}_1)ln\left(\frac{\hat{s}_1-x}{\hat{s}_s}\right)+x^2(\tilde{n}_s+x)\right]}{(x+\tilde{n}_s)^2}}_{\text{Much smaller than 1}}\right] > 0.$$

As such, the family DUH satisfies PT, in addition to GSYM, INV, IND, and PPT. To see that the family DUH satisfies CON, see that $\hat{P}_1$ and all the $\tilde{P}_g$'s are invariant to the addition of zero-groups, so I only need to check that ψ_p decreases when G increases by 1. In other words, I only need to show that $\left(\sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}$ increases in G. First, observe the following inequality:

$$\sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G-1}\right|^p = \sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G}+\frac{1}{G}-\frac{1}{G-1}\right|^p$$
$$\leq \sum_{g=2}^{G}\left(\left|\tilde{P}_g-\frac{1}{G}\right|^p+\left|\frac{-1}{G(G-1)}\right|^p\right) = \sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G}\right|^p+\frac{G-1}{G^p(G-1)^p}$$
$$< \sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G}\right|^p+\frac{1}{G^p} = \sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G}\right|^p+\left|0-\frac{1}{G}\right|^p = \sum_{g=2}^{G+1}\left|\tilde{P}_g-\frac{1}{G}\right|^p$$

Taking the p^{th} root on both side, we have

$$\left(\sum_{g=2}^{G}\left|\tilde{P}_g-\frac{1}{G-1}\right|^p\right)^{\frac{1}{p}} < \left(\sum_{g=2}^{G+1}\left|\tilde{P}_g-\frac{1}{G}\right|^p\right)^{\frac{1}{p}}. \quad \square$$

To see that the family DUH satisfies UNI, verify that $\Phi(1,0,\ldots,0)=0$ and $\Phi\left(\frac{1}{G},\ldots,\frac{1}{G}\right)=1$, for all $G \in \mathbb{N}\setminus\{1,2\}$. $\square$

Appendix C Characterization Axioms of HHI and SE

Axiom C.1 (Expandability (EXP))**.** *Take an arbitrary system $s = (s_1, \ldots, s_G)$ of G groups. Φ satisfies expandability if*

$$\Phi(s_1, \ldots, s_G) = \Phi(s_1, \ldots, s_G, 0).$$

It should be clear that neither EXP nor CON attempts to pin down the functional form of a meaure. Rather, these two opposing axioms serve as the divide between a unit for concentration and a unit for heterogeneity.

As discussed extensively by Atkinson (1970), the fact that PDT only induces partial ordering implies that specific functional forms can always be chosen to induce different total orders when neither system's distribution second-order stochastically dominates the other.[11] The functional forms of both HHI and SE were able to be uniquely characterized because of this feature.[12] HHI uses EXP and the replication principle (REP), and SE uses EXP and Shannon's additivity (SADD).

Axiom C.2 (Replication Principle (REP))**.** *Take an arbitrary system s such that $|s| = G$. Φ satisfies the replication principle for concentration if replicating a system k times divides the system concentration by k.*

For example, take $k \in \mathbb{N}$,

$$\frac{1}{k}\Phi(s_1, \ldots, s_G) = \Phi\left(\underbrace{\frac{s_1}{k}, \frac{s_1}{k}, \ldots, \frac{s_1}{k}}_{\text{Sum to } s_1}, \frac{s_2}{k}, \frac{s_2}{k}, \ldots, \underbrace{\frac{s_G}{k}, \ldots, \frac{s_G}{k}}_{\text{Sum to } s_G}\right).$$

REP pins down the cardinal meaning of the unit by linking the multiplication of the unit to how many times a system is divided/replicated into a system with more groups. Chakravarty and Eichhorn (1991) show that if a concentration unit C can be represented as a self-weighted quasilinear mean, then C is the Hannah-Kay index of concentration if and only if C satisfies the fundamental axioms and REP.[13] Since HHI is $H^n_{\alpha=2}(s)$, it is the unique self-weighted quasilinear concentration unit that satisfies the fundamental axioms, EXP, and REP. Notice that REP applies only when *every group* is broken into multiple groups. To prescribe how the index behaves when only one group is broken up, SE forgoes REP and restricts the measure based on Shannon's additivity instead.

[11] Newbery (1970) shows that no additive functional forms can be chosen to induce the same order as the Gini coefficient. This impossibility theorem stems from and reaffirms the point made by Atkinson (1970).

[12] Similarly, I use this feature to uniquely characterize the descriptive units of heterogeneity.

[13] A relative concentration index $C : D \to \mathbb{R}$ is called a self-weighted quasilinear mean if for all $n \in \mathbb{N}$, $x \in D^n$, $C^n(x)$ is of the form

$$C^n(x) = \phi^{-1}\left[\sum_{g=1}^{G} P_g \phi\left(P_g\right)\right],$$

where $\phi : (0, 1] \to \mathbb{R}$ is strictly monotonic.

Axiom C.3 (Shannon's additivity (SADD)). *Let $s = (s_1, \ldots, s_G)$ be anarbitrary system. Define $s_{gj} \geq 0$ such that $s_g = \sum_{j=1}^{m_g} s_{gj}$,*
$\forall g \in \{1, \ldots, G\}$ and $\forall j \in \{1, \ldots, m_g\}$. The measure Φ satisfies Shannon's additivity if

$$\Phi(s_{11}, \ldots, s_{Gm_G}) = \Phi(s_1, \ldots, s_G) + \sum_{g=1}^{G} \frac{s_g}{||s||_1} \cdot \Phi\left(\frac{s_{g1}}{s_g}, \ldots, \frac{s_{gm_g}}{s_g}\right).$$

By setting $m_g = 1$, $\forall g \in \{1, \ldots, G-1\}$ and $s_{G'} = s_G + s_{G+1}$), SADD implies htat

$$\Phi(s_1, \ldots, s_G, s_{G+1}) = \Phi(n_1, \ldots, s_{G'}) + \frac{s_{G'}}{s_1 + \ldots + s_{G-1} + s_{G'}} \cdot \Phi\left(\frac{s_G}{s_{G'}}, \frac{s_{G+1}}{s_{G'}}\right).$$

SADD pins down how the decomposition of group(s) in a system should influence the unit.[14]

For detailed proofs of the unique characterization of SE and explanations of SADD, readers should refer to Suyari (2004) and Chakrabarti, Chakrabarty et al. (2005).

[14] An example of decomposing a group is to split the sales of Macs into Mac desktops and Mac laptops, for the purpose of measuring the heterogeneity of Apple's revenue streams.