Descriptive Units of Heterogeneity: An Axiomatic Approach to Measuring Heterogeneity

Willy Chen*

2024

Abstract

This paper addresses the challenge of measuring between-group heterogeneity in systems, as existing measures lack cardinal interpretation and comparability across systems/population. Using an axiomatic approach, I highlight the strengths and limitations of existing measures and generalize properties that need to be satisfied by better alternatives. Using these axioms, I propose the class of measures called the *Descriptive Units of Heterogeneity* (DUH), a hybrid solution to prior limitations without limiting the applicable contexts. DUH achieves the generalized comparability of concentration units while still able to reflect changes in the distribution of small groups in the population. Hence, DUH provides a valuable tool for empirical researchers studying heterogeneity in systems in various contexts, such as racial composition in a city or revenue shares by products of a firm.

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

*Michigan State University: willyc@msu.edu

1 Finding a New Way to Measure Heterogeneity

My objective is to measure the heterogeneity present in a system – a collection of distinct groups with more than one elements. For example, consider the revenue streams by each Apple product a group, the sales number for each product is population/number of elements, in that group. The collection of these groups—iPhone, iPad, Mac, Services, and Wearables & Accessories—in the system, Apple's revenue streams. Heterogeneity in a system means that there is variation in the outcome of interest, while homogeneity is the lack of variation. For example, if iPhone sales is 90% of Apple's revenue, then the revenue streams of Apple is less heterogeneous than if iPhone sales and iPad sales are both 45% of Apple's revenue.

Measuring heterogeneity is, at its heart, a dimension-reduction problem. When studying a complex population, one hopes to simplify the complexity without losing the big picture. Deciding the essential elements for *describing* a system is the key to defining a measure that can tractably identify changes in a system. I pin down perfect heterogeneity as when all groups in the system has the same number of elements, and perfect homogeneity as when all but one group in the system has non-zero amount of elements. The next task is to balance the influence of large groups and the influence of the distribution of elements across groups.

The heterogeneity of a system is commonly measured in one of two ways: Dispersion units and Concentration units. Dispersion units focus on measuring the distance between the observed population distribution to a benchmark distribution, yielding concise interpretations at the cost of comparability between systems. Atkinson et al. (1970) showed that comparing any two systems using the distance between distribution requires significant restrictions on the domain of comparable systems; otherwise, different measures can be made to rank any two systems in opposing ways. Within the bounds of said restrictions, the interpretation of any dispersion unit is simply "The higher the number, the higher the heterogeneity." One well-known example of a dispersion unit is the *Gini coefficient*. While its simple interpretation contributed to its popularity, Schwartz and Winship (1980) showed that many empirical researchers fail to check for said restrictions when using the *Gini coefficient* to rank countries by income inequality and produced results that are at times contradictory.

Concentration units focus on measuring the richness of information from select subgroups of the population. These units differ from dispersion units by having generalized compatibility between systems, but they understate the information provided by small subgroups. By emphasizing the influence of large groups, concentration units reflect changes in heterogeneity in systems well when a large group shrinks. However, this feature results in concentration units producing negligible changes in heterogeneity when there are drastic changes between the small groups. Such feature is not a desirable property for a measure of heterogeneity.

In this paper, I propose a new yet intuitive way to think about the make-up of heterogeneity by separating it into *contribution from the relative size of the largest group* (majority) and *contribution from the evenness of the rest of the groups* (minority) and show that this, in fact, builds on the existing paradigm.

Related Literature. This paper is not the first attempt in this strand of literature at an axiomatic characterization measure. In fact, there is abundant existing literature that axiomatize Gini coefficient (Schwartz and Winship 1980; James and Taeuber 1985), *Herfindahl–Hirschman Index* (HHI) (Kvålseth 2022; Chakravarty and Eichhorn 1991), and *Shannon's Entropy* (SE) (Nambiar et al. 1992; Suyari 2004; Chakrabarti et al. 2005). I call the axioms that are common the *Fundamental axioms*, and the ones that uniquely characterize each measure the *Characterization axioms*.

Fundamental axioms are axioms that all measures of heterogeneity should satisfy. These axioms pin down when two systems of the same number of groups are equally heterogeneous and how they should be ordered when they are marginally different. Existing literature showed that while Gini, HHI and SE satisfy the fundamental axioms; further, HHI and SE are uniquely characterized by their own sets of characterization axioms. The characterization axioms generally but not necessarily pin down cardinal interpretations for each measure. I define new characterization axioms and show that these axioms keep the generalized comparability between systems and provide cardinal interpretation to the measure. My new axioms, along with the fundamental axioms, characterize a class of indices that focus on making the comparisons between system *descriptive*—generally comparable, cardinally interpretable, and reflective of small group changes. This class of indices is termed the *Descriptive Units of Heterogeneity*.

I put forth the idea of a *reasonable universe of groups* for practical uses of any measure of heterogeneity. The reasonable universe of groups is the set of grouping labels that the researcher deems reasonable and comparable. I use changes in racial composition in San Francisco and changes in Apple's revenue shares as empirical examples to illustrate the strength of my index in accounting for both the contribution of the majority groups and the evenness of the minority groups, a feature and a curse of the concentration units by design.

The rest of the paper proceeds as follows. Section 2 reviews existing measures in detail and provides motivation for a new paradigm using features of select existing units of heterogeneity. Section 3 lists and discusses the *Fundamental Axioms* and the *Characterization Axioms* along with my proposed new paradigm. Section 4 defines the *Descriptive Unit of Heterogeneity* and compares this new index to the existing concentration units. Section 5 presents empirical examples to highlight the strength of the descriptive unit of heterogeneity in various settings.

The paper concludes by outlining the distinctions between the descriptive units of heterogeneity and established indices, while also elucidating best practices for its application. The appendix is utilized to provide the requisite proofs that establish the uniqueness of my index.

2 Lessons from Existing Units of Heterogeneity

First, I define the primitives of the paradigm I use to discuss heterogeneity. Let $\Theta = \{\theta_1, \dots, \theta_G\}$ be a **universe** of $G \in \mathbb{N}$ distinct groups/categories. A system S is a mapping from Θ to $\mathbb{Z}_+^G$ such that $S = (n_1, \ldots, n_g \ldots, n_G)$ is a $1 \times G$ vector where n_g is a positive integer that represents the number of elements in the group θ_g. A system S with population n_S is thus the collection of groups θ_g each with n_g elements.

The measure of heterogeneity is then a mapping $\Phi : \mathbb{Z}_+^G \to \mathbb{R}$ such that for any two systems S and S'

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

For example, let S be Michigan State University, S maps the universe of groups $\Theta = \{faculty, staff, students\}$ to the number of faculty n_{fac}, staff n_{sta}, and students n_{stu} at Michigan State. Heterogeneity in this system is the presence of mixture, e.g. the presence of a mix of faculty, staff, and students. Homogeneity is the lack of mixture, meaning only one or two of these groups are present in the system.

Definition 1: A system S of G groups is said to achieve **maximum heterogeneity** if it can be represented as a scalar multiple of the identity vector of size $G \in \mathbb{N}$:

$$S = (\underbrace{n, n, \ldots, n}_{\substack{G \text{ groups each} \\ \text{with } n \text{ elements}}}) = n \cdot (1, 1, \ldots, 1)$$

Definition 2: A system S' of G groups is said to achieve **minimum heterogeneity/perfect homogeneity** if it can be represented as a $1 \times G$ vector where all but one entry are 0:

$$S' = (0, 0, \ldots, 0, n, 0, \ldots, 0) = n \cdot (0, 0, \ldots, 0, 1, 0, \ldots, 0)$$

The one-dimensional (presence of mixture) nature of this definition makes it convenient for any measure to be bounded between $\Phi(S) \in \mathbb{R}_{++}$ and $\Phi(S) = 0$. The harmlessness of this generalization is evidenced by existing measures of heterogeneity, even when they are constructed with different goals in mind. These units can be generally separated into two categories—Dispersion units and Concentration units.

This section provides brief discussions of select dispersion and concentration units that motivate the formalization of axioms in section 3. I will show that even though dispersion units often have limited use cases, the direct inference they enable make them an attractive option. On the other hand, concentration units excels at generalizing comparisons between systems while they focus on indirect inferences limited to the majority in a system. The axioms discussed in section 3 thus are the products of my desire to find an index that can have the strength of both types of units while suffering minimally from potential limitations.

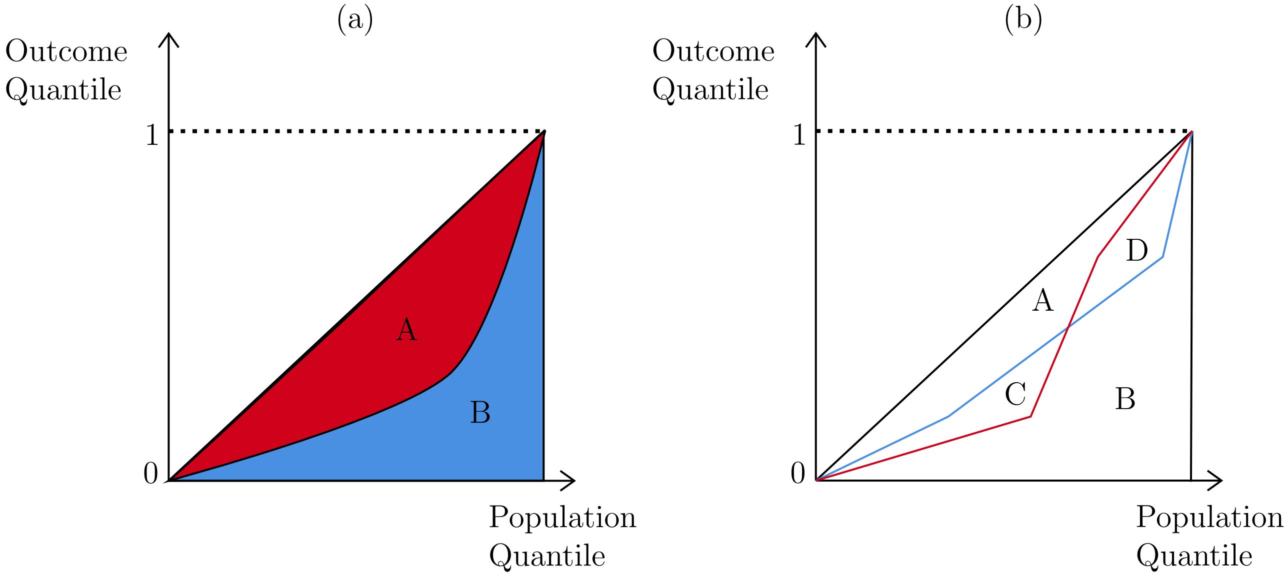
2.1 Dispersion Units (Gini Coefficient and Dissimilarity Index)

Readers familiar with the economics of inequality literature are likely familiar with the Gini coefficient or the Dissimilarity Index, both of which are dispersion units.

The most commonly used form of the *Gini coefficient* maps the percentiles of a single ordered outcome to the percentiles of the population. It compares the distribution of said outcome to the uniform distribution. The idea is simple, but powerful. Because the outcome is standardized, it is not hard to understand the sentence "the top x% of households in the US earns the top y% of income." In a uniformly distributed world, x and y should be equal, and if it is not, then there is inequality.

Part (a) in Figure 1 is how this is taught in most undergraduate development classes. The Gini coefficient of an economy is defined as $\frac{A}{A+B}$, and higher A means higher inequality. However, this simple interpretation comes at the cost of a restrictive assumption—the *Lorenz criterion*. The Lorenz criterion states that using the Gini coefficients of two systems to compare inequality implicitly assumes that the Lorenz curves of the two systems do not cross (Atkinson et al. 1970; Marshall et al. 1979; Schwartz and Winship 1980; James and Taeuber 1985).

Figure 1: Gini Coefficient: Definition and Comparison



For example, if we look at part (b) of Figure 1, we can compare a country with $B + C + D$ under the Lorenz curve to a country with just B under the Lorenz curve, and say that the second country has more inequality (because A is smaller than $A + C + D$). But we cannot use the Gini coefficient of a country with $B+C$ under the Lorenz curve to the Gini coefficient of a country with $B + D$ under the Lorenz curve, and say either country has more inequality than the other.Schwartz and Winship (1980) provides several examples explaining why the Lorenz criteria significantly limits use cases of the Gini coefficient. They also show that the Lorenz criterion is often neglected in empirical research, leading to contradictory results at times. I omit this discussion as that is not in the scope of this paper.

On the other hand, *Dissimilarity index*, a popular unit of sorting in the ineuqality literature, does not have implicit restrictions like the Lorenz criterion. The idea of dissimilarity index is simple, yet elegant—Comparing the distribution of *two groups* in a neighborhood to the overall distribution of those same groups in the city reveals sorting, i.e., dissimilarity between the neighborhoods using random assignment as the benchmark. By summing up the differences between each neighborhood's observed distribution and its counterfactual distribution without sorting, one can infer higher dissimilarity from just a higher sum. For example, the dissimilarity index for income sorting in a city with $I \in \mathbb{N}$ neighborhoods is:

$$DI = \frac{1}{2}\sum_{i=1}^{I}\left|\frac{\#LowIncome_i}{\sum_{i=1}^{I}\#LowIncome_i} - \frac{\#NotLowIncome_i}{\sum_{i=1}^{I}\#NotLowIncome_i}\right|.$$

This index has been useful in the broad literature of income/racial sorting, but it does suffer from one design flaw—there can only be two groups. It is straightforward to interpret an increase in DI when it is made up of differences in proportions of a binary outcome, but the nature of differences between proportions obscures any natural extension to systems of three or more groups.

These are not the only dispersion units[1], but they are excellent examples of indices that are useful because of their straightforward interpretation, despite the limitations discussed. Overall, comparing the observed distribution to a benchmark distribution is intuitive and easy to understand. But their limitations make them less practical in modern empirical research where one may need general comparison between systems. Concentration units, on the other hand, suffers from just the opposite.

2.2 Concentration Units (Herfindahl–Hirschman Index and Shannon's Entropy)

A simple fix to gain generalized comparability across systems is to use the Concentration units to measure heterogeneity. Let there be a system S with G groups. The *Herfindahl–Hirschman Index* (HHI) and its complement Gini-Simpson Index of system S are defined as:

$$HHI(S) = \sum_{g=1}^{G}\left(\frac{n_g}{n_S}\right)^2, \quad GSI(S) = 1 - HHI(S).$$

HHI, and subsequently GSI, has an alluringly simple interpretation—The probability of 2 random draws with replacement, from the system S, being from the same group in S (in this case, GSI is quite literally the complement of that). The elegance of this measure had historically outshined its flaws, perhaps because it is developed for contexts where such flaws were intentional.

HHI disproportionately accounts for changes in large groups. Observe that, due to the squaring of group proportions, any group accounting for less than, say 10%, of the system

[1] Nunes et al. (2020) discusses several other dispersion units in the context of ecology, biology, and medicine.

has fewer than 1% impact on the system's HHI. For example, say that MSU in 2020 has 80% students, 15% staff, and 5% faculty, and that those respective numbers change to 80%, 19%, and 1% in 2030. The respective HHIs are 0.665 and 0.6762. In this case, the faculty population shrank by 80%, but the changes in HHI is barely noticeable.

Conventional utilization of HHI primarily pertains to the assessment of market power of a firm as a function of its market share. Consequently, the significance of an 80% reduction in a firm's presence, initially constituting a mere 5% of the market, may appear negligible if the 4% were redistributed to another minor competitor. This necessitates a contextual confinement of HHI's application to discussions concerning market shares and concentrations, rather than employing it as a measure of heterogeneity.

Moreover, the simple interpretation of HHI can obscure obvious differences between systems; two systems, $(48\%, 48\%, 4\%)$ and $(60\%, 30\%, 10\%)$ will yield 0.4624 and 0.46 HHI. The heterogeneity of these systems are reasonably different, but that is hardly reflected in the difference in their HHIs. Unlike the dispersion units discussed above, when a system's HHI increases by some number x, the only interpretation is the probabilistic one, which is seldom descriptive in the case of measuring heterogeneity. Combining with the fact that it accounts mostly for the larger groups, using HHI costs researchers much information about movements within the smaller groups.

Also suffering from this flaw is the commonly used measure of uncertainty in information theory—Shannon's Entropy, defined as:

$$SE(S) = -\sum_{g=1}^{G} \left[\frac{n_g}{n_S} \cdot ln \left(\frac{n_g}{n_S} \right) \right].$$

One key thing to notice is that in both of these units, zero-groups, groups with zero elements ($n_g = 0$), do not affect the measure at all. This property is intuitive when used to capture market shares or uncertainty, but it should not be salient when used to measure heterogeneity. One likely does not care about whether Starbucks has any market power in the pencil lead industry, and one need not care about whether the probability of finding pencil lead in their coffee is accounted for or not. But why should the same be said about any measure of heterogeneity?

Comparability between systems hinges on what the comparison is to capture and whether two systems are similar enough. Comparing apples to oranges might not necessarily be nonsensical like the popular idiom suggests. In the proper context, I can compare an apple to an orange on density, brightness of color, amount of sugar per ml of water, etc. I can even compare an apple to an orange on how good each fruit is at being a citrus (clearly, the apple will lose). What I cannot do is say that an apple is x times denser than an ice cube than an orange is y times rounder than a bowling ball.

To effectively compare heterogeneity between two systems, it is crucial to establish the qualifiers that make them comparable. When assessing system complexity, considering all groups, including those with zero elements, provides a baseline for measurement. Thus, comparing systems with varying group counts requires viewing the system with fewer groups as encompassing the additional groups with zero elements, ensuring an accurate evaluation of heterogeneity. From Chakravarty and Eichhorn (1991):

> [Zero-groups having no impact on the measure] means that the index value does not alter if there is addition or deletion of a firm with zero output. This particular principle shows the fundamental difference between indices of concentration and indices of inequality. It is generally agreed that adding (deleting) an individual with zero income to (from) a population increases (decreases) inequality. But firms which produce no output should not have any impact on concentration. (p.104)

Catering an index to the inclusion of zero-groups is hardly revolutionary. In fact, a normalized version of HHI attempts to solve that issue by revising the formula to:

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

This index improves system compatibility by accounting for zero-groups via the normalization, but it is done at the cost of HHI's simple probabilistic interpretation. Consider the following two systems S and S':

$$S = (0.4, 0.4, 0.2)$$
$$S' = (0.5, 0.3, 0.1, 0.1)$$

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 the comparison of these two NHHIs is unlikely to be convincing. Once we account for zero-groups and 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.

2.3 The Shopping List for An Unknown Recipe

Learning from the strengths and weaknesses of these units, I want to create an index that can yield simple interpretations when comparing any two systems without giving up too much for complexity. The following list are desired features that either directly or indirectly motivate the axioms that my index Φ should satisfy:

- A $\Phi(S') = x \cdot \Phi(S)$ can be interpreted as S' is x times as heterogeneous as S via some channel.
- Φ and existing units/indices share some basic properties.
- Φ accounts for the presence of zero-groups.

With these in mind, I will discuss the axiomatization of units of heterogeneity.

3 Axioms for Units of Heterogeneity

Axioms from existing literature can be generally separated into two categories:

1. *Fundamental Axioms:* Axioms that all measures of heterogeneity should satisfy. These axioms pin down when two systems are equally heterogeneous and how two marginally different systems should be ordered.

2. *Characterization Axioms:* Axioms that uniquely characterize measures by imposing cardinal interpretation.

3.1 Fundamental Axioms

[SYM] Group Symmetry. Given the same number of groups (G) in the system, the heterogeneity of a system is invariant to relabelling groups.

$\forall n_a, n_b, n_c \in \mathbb{N}$,

G. \ S.	S1		S2		S3
A	n_a		n_b		n_c
B	n_b	$\sim$	n_a	$\sim$	n_b
C	n_c		n_c		n_a

SYM is an intuitive axiom as it enables the index to focus on the distribution over groups in a system, rather than sizes of individual groups. Satisfying *SYM* means any two systems with the same number of groups can be compared. By focusing on distribution over the same number of groups, *SYM* enables comparisons between systems mapping different universes of groups, so long as the two universe have the same number of groups.

In a similar generalization effort, any index of heterogeneity should focus on the proportion (relative sizes) of each group and not the absolute sizes of each group. This generalization yields this next axiom - *Scale Invariance.*

[INV] Scale Invariance. Given two systems with the same proportions in each group, the heterogeneity of the combined system must be the same. Without loss of generality, this can be generalized to the following:

$\forall n_a, n_b, n_c, \in \mathbb{N},\ \lambda \in \mathbb{R}_{++}$,

G. \ S.	S1		S2		S1+S2
A	n_a		λn_a		$(1+\lambda)n_a$
B	n_b	$\sim$	λn_b	$\sim$	$(1+\lambda)n_b$
C	n_c		λn_c		$(1+\lambda)n_c$

INV allows for a further generalization of the systems to have groups with sizes of any positive real numbers. This axiom ensures that the index reflects only the distribution of sizes in a system rather than the absolute sizes of the groups.

The first two axioms give us a convenient way to equate heterogeneity between systems with different sizes but the same distribution. Having generally pin down heterogeneity equivalence between systems, my next task is to pin down the ordering of marginally different systems.

[PDT] Principle of Diminishing Transfers. *Holding the order of groups constant*, transferring population from a larger group to a smaller group increases heterogeneity. The increase increases in the difference between the two groups.

Take any $n_a, n_b, n_c, n_d \in \mathbb{N}$ such that $n_a > n_b > n_c > n_d$ and let $\varepsilon < \min\left\{n_c - n_d,\, n_b - n_c,\, \dfrac{n_a - n_b}{2}\right\}$, then

$$
\begin{array}{|c|}\hline \text{G.} \backslash \text{S.} \\ \hline \text{A} \\ \hline \text{B} \\ \hline \text{C} \\ \hline \text{D} \\ \hline \end{array}
\quad
\begin{array}{|c|}\hline \text{S1} \\ \hline n_a \\ \hline n_b \\ \hline n_c \\ \hline n_d \\ \hline \end{array}
\prec
\begin{array}{|c|}\hline \text{S2} \\ \hline n_a - \varepsilon \\ \hline n_b + \varepsilon \\ \hline n_c \\ \hline n_d \\ \hline \end{array}
\prec
\begin{array}{|c|}\hline \text{S3} \\ \hline n_a - \varepsilon \\ \hline n_b \\ \hline n_c + \varepsilon \\ \hline n_d \\ \hline \end{array}
\prec
\begin{array}{|c|}\hline \text{S4} \\ \hline n_a - \varepsilon \\ \hline n_b \\ \hline n_c \\ \hline n_d + \varepsilon \\ \hline \end{array}
$$

This axiom originated as the *Principle of Transfers*, first formulated by Dalton (1920), "...if there are only two income-receivers, and a transfer of income takes place from the richer to the poorer, inequality is diminished" (p.351). Over time, scholars have discussed whether the *decrease* in inequality is constant, increasing, or decreasing, in the difference in income between the richer and the poorer. For a measure of heterogeneity, I believe that the decrease in inequality should be increasing in the difference of proportions of the two groups. Therefore, it is only reasonable that *PDT* is the third and last fundamental axiom.

When comparing distributions using quantiles, the most common use case, assuming *SYM*, *INV*, and *PDT*, is equivalent to assuming the *Lorenz Criterion*—the Lorenz curves of two distributions do not cross. However, the general case of the *Lorenz Criterion* is only equivalent to *PDT*. Equivalently, intuitive as *PDT* may seem, it still only yield partial ordering of distributions (Rothschild and Stiglitz 1969; Atkinson et al. 1970; Rothschild and Stiglitz 1973; Kolm 1976). The implementation of *PDT* in a measure of inequality is well discussed in mathematics and statistics as *majorization*. Chapter 1 of Marshall et al. (1979) has an excellent discussion on the limits of using majorization to study inequality, and how further assumptions/axioms are necessary to propel the indices of inequality in practical use cases.

3.2 Characterization Axioms

A New Way to Think About Heterogeneity—Before I continue in the world of axiomatization, I want to propose a new way to think about heterogeneity. All the units I have discussed up to this point treat each group in the system numerically equally, partially in order to ensure *SYM*. However, group symmetry can be satisfied without each group getting identical numerical treatment. Rather, it can be satisfied simply by giving the same *types* of groups the same numerical treatment.

Notice that for a index that can be used for generalized comparison between systems, it needs to be able to order the heterogeneity of systems of any positive integer number of groups. Notice that for any system, there is always the largest group and the remaining groups. What I propose is an index that treats this largest group differently than the rest of the groups in the system. The advantage to this new way of thinking about heterogeneity is that if the influence of the largest group on the index can be orthogonal to the influence of the rest of the groups, we can equate any changes in the one-dimensional index with the equivalent change(s) in either group(s) while holding the other constant.

By *SYM*, groups in each system can be ordered by the size of each group, meaning that giving the largest group a different treatment than the rest of the groups does not violate group symmetry. For convenience, I will now call the largest group the *majority group* and the rest of the groups the *minority groups*.

Under this new paradigm, I can refine the definition of a unit of heterogeneity. Recall our earlier definition—The heterogeneity of system S is measured by $\Phi : S \to \mathbb{R}_+$ such that

$$\Phi(n_1, \ldots, n_G) \geq \Phi(n'_1, \ldots, n'_G) \iff S \text{ is weakly more heterogeneous than } S'$$

Now consider the paradigm where system heterogeneity is comprised of into two parts:

1. Relative size of the Majority Group: $P_1 = \frac{n_1}{n_1+\cdots+n_G}$
2. Relative size(s) of the Minority Group(s): $P_2, \ldots, P_G$

An index for heterogeneity is then $\Phi = \Phi(\varphi, \psi)$ where φ is the influence of the relative sizes of the majority groups and ψ the influence of the relative sizes of the minority groups. If we can find well-defined functions ϕ and ψ, we can then order any two systems that we deem comparable in the context of our work. The rest of the axioms will focus on refinements of *PDT* for gaining complete ordering of system heterogeneity in an intuitive way.

For the influence of P_1 and P_2 through P_G to be orthogonal, consider the *Independence Axiom.*

[IND] Independence. The influence of the relative size of the majority population on the unit of heterogeneity should be independent of the relative sizes of the minority groups, and vice versa.

$$\varphi(n_1, n_2, \ldots, n_G) = \varphi\left(\frac{n_1}{n_1 + \cdots + n_G}\right) = \varphi(P_1).$$
$$\psi(n_1, n_2, \ldots, n_G) = \psi(n_2, n_3, \ldots, n_G) = \psi\left(\frac{n_2}{n_2 + \cdots + n_G}, \ldots, \frac{n_G}{n_2 + \cdots + n_G}\right).$$

To satisfy *IND*, we need a function that omits P_1 by design and does not violate any of the previous axioms. Let $\tilde{P}_g = \frac{n_g}{n_2+\cdots+n_G}$, $\forall g \in \{2, \ldots, G\}$,

$$\tilde{P}_g = \frac{n_g}{n_2 + \cdots + n_G} = \frac{\frac{n_g}{n_1+\cdots+n_G}}{\frac{n_2}{n_1+\cdots+n_G} + \cdots + \frac{n_G}{n_1+\cdots+n_G}} = \frac{P_g}{P_2 + \cdots + P_G}.$$

The immediate implication of *IND* is that the unit Φ must be functionally separable, and the definition using relative group sizes implies *INV*.

Following the similar methodology in dispersion units for easy interpretations, I want ψ to reflect the evenness in the minority groups distribution. Specifically, I define ψ to be the distance between the observed distribution in the minority groups and the ideal uniform distribution in the minority groups. This means that ψ can be characterized as a class of functions by all finite p-metric d_p in $\mathbb{R}^n$.

Definition 3: A function $\psi : \mathbb{R}_+^{G-1} \to \mathbb{R}_+$ is a measure of evenness in minority group distribution if it is of the following form:

$$p \in \mathbb{R}_+,\ \psi_p(S) = \psi\left(\tilde{P}_2, \ldots, \tilde{P}_G\right) = 1 - \left(\sum_{g=2}^{G} \left|\tilde{P}_g - \frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}.$$

Proposition 1: Consider an index $\Phi_p = \Phi(\varphi, \psi)$ that satisfies *SYM*, *INV*, and *IND*. Holding P_1 constant[2], if Φ is strictly increasing in ψ_p, then $\Phi(\varphi(P_1), \psi)$ satisfies the *Principle of Diminishing Transfers* if and only if $p > 1$.

Sketch of Proof: Given the absolute value function, the proof simply follows the mechanics of Jensen's inequality. The full proof is included in the appendix.

Having formally defined ψ_p, I want to pin down ϕ with an axiom that fully describes changes in heterogeneity when ψ is fixed. In pursuit of an index whose changes are easy to interpret, I propose a minimalist refinement of *PDT* that builds on the notion of *IND* - *Principle of*

[2]This proposition suggests that we need to be careful with the functional form of $\varphi(P_1)$. To satisfy *PDT*, it must be that when there is a transfer from the majority group to the minority group, the increase in $\varphi(P_1)$ must dominate the decrease in ψ_p in the case where evenness decreases.

Proportional Transfers.

Remark: It is simple to verify that a higher p gives more weight to groups with proportions that are farther from $\frac{1}{G-1}$. For the purpose of reflecting more information from the evenness of minority groups, I recommend using d_2, the Euclidean distance, resulting in ψ_2 as the measure of evenness.

[PPT] Principle of Proportional Transfers. *Holding the order of groups constant*, a transfer from the majority group proportionally to the minority groups that reduces P_1 to P_1^α increases heterogeneity by a factor of α.

G. \ S.	S1		S2
θ_1	P_1		P_1^α
θ_2	P_2	$\sim$	$P_2 + \frac{P_2}{P_2+P_3}(P_1 - P_1^\alpha)$
θ_3	P_3		$P_3 + \frac{P_3}{P_2+P_3}(P_1 - P_1^\alpha)$

Then

$$\Phi(S_1) < \alpha\Phi(S1) = \Phi(S2).$$

PPT gives changes in the index a simple interpretation. If $2\Phi(S1) = \Phi(S2)$, then one can say that $S2$ *is twice as diverse as* $S1$ because it has the equivalent heterogeneity as if the majority group proportion in $S1$ shrunk by the power of 2 while still being the majority group, holding the same evenness in the minority groups.

Thus far, the axioms are laid out for cases when two systems have the same number of groups. In fact, existing literature uniquely characterize concentration units through how the measures behave when the number of groups change and the cardinal interpretation that follows. This next set of axioms pin down the changes to an unit when a group (with zero elements or otherwise) is added to the system. The first axiom of note is the axiom of *Expandibility*, an axiom that is favored by concentration units.

[EXP] Expandability. $\Phi(n_1, \ldots, n_G)$ satisfies *Expandability* if

$$\Phi(n_1, \ldots, n_G) = \Phi(n_1, \ldots, n_G, 0).$$

As discussed in Chakravarty and Eichhorn (1991), *EXP* is a salient reason why concentration units such as HHI and SE should not be used to measure inequality. This axiom is reasonable as the information provided by the presence of zero-groups is not pertinent when measuring concentration. Nevertheless, it would be unreasonable to say the same for a unit meant to measure the inequality in or the heterogeneity of a system.

As a contrarian axiom to *EXP*, I propose the axiom of *Contractibility*.

[CON] Contractibility. Φ satisfies *Contractibility* if adding one 0-group to a system of G groups decreases heterogeneity of the system.

$$\Phi(n_2, \ldots, n_G, 0) < \Phi(n_2, \ldots, n_G).$$

A practical implication of *CON* is that the comparison between systems with a unit assumes that the two systems have the same number of groups, even if some groups have 0 elements. It should be clear that neither *EXP* nor *CON* attempts to pin down the functional form of a unit. Rather, these two opposing axioms serve as the divide between a unit for concentration and a unit for heterogeneity.

As discussed extensively in Atkinson et al. (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 dominates the other. Both HHI and SE[3], use this feature to motivate axioms for functional forms to uniquely characterize the measure. HHI uses *EXP* and the *Replication Principle* (REP) and SE uses *EXP* and *Shannon's Additivity* (SADD).

REP pins down the cardinal meaning of the unit by linking the multiplication of the unit to the how many times a system is divided/replicated into a system with more groups.

[REP] Replication Principle. $\Phi(n_1, \ldots, n_G)$ satisfies the *Replication Principle* for concentration if replicating a system k times divides the system concentration by k.

$\forall k \in \mathbb{N}$,

$$\frac{1}{k}\Phi(n_1, \ldots, n_G) = \Phi\left(\underbrace{\frac{n_1}{k}, \frac{n_1}{k}, \ldots, \frac{n_1}{k}}_{\text{Sum to } n_1}, \frac{n_2}{k}, \frac{n_2}{k}, \ldots, \underbrace{\frac{n_G}{k}, \ldots, \frac{n_G}{k}}_{\text{Sum to } n_G}\right).$$

Chakravarty and Eichhorn (1991) and Schwartz and Winship (1980) show that any concentration unit satisfying *SYM*, *INV*, and *PDT* satisfies the *Lorenz Criterion.* Once such concentration unit is the *Hannah-Kay* class of concentration unit, with perception α and $n \in \mathbb{N}$ firms in industry $S \in D^n$, defined as:

$$H_\alpha^n(S) = \begin{cases} \left[\sum_{i=1}^{n} s_i^\alpha\right]^{\frac{1}{\alpha-1}} & \text{if } \alpha > 0,\ \alpha \neq 1 \\ \prod_{i=1}^{n} s_i^{s_i} & \text{if } \alpha = 1 \end{cases}.$$

Chakravarty and Eichhorn (1991) futher shows that a concentration unit C can be repre-

[3]and later on, the *Descriptive Units of Heterogeneity*

sented as a *self-weighted quasilinear mean*[4]. Then C is the Hannah-Kay index of concentration if and only if C satisfies the replication principle. Notice that HHI is $H^n_{\alpha=2}(S)$.

SADD pins down how decompositions[5] of group(s) in a system should influence the unit.

[SADD] Shannon's Additivity. Define $n_{gj} \geq 0$ such that $n_g = \sum_{j=1}^{m_g} n_{gj}$, $\forall g \in \{1, \dots, G\}$, $\forall j \in \{1, \dots, m_g\}$

$\Phi(n_1, \dots, n_G)$ satisfies *Shannon's Additivity* if

$$\Phi(n_{11}, \dots, n_{Gm_G}) = \Phi(n_1, \dots, n_G) + \sum_{g=1}^{G} \frac{n_g}{n_S} \cdot \Phi\left(\frac{n_{g1}}{n_g}, \dots, \frac{n_{gm_g}}{n_g}\right).$$

which implies (by setting $m_g = 1$, $\forall g \in \{1, \dots, G-1\}$ and $n_{G'} = n_G + n_{G+1}$),

$$\Phi(n_1, \dots, n_G, n_{G+1}) = \Phi(n_1, \dots, n_{G'}) + \frac{n_{G'}}{n_1 + \ldots + n_{G-1} + n_{G'}} \cdot \Phi\left(\frac{n_G}{n_{G'}}, \frac{n_{G+1}}{n_{G'}}\right).$$

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

Remark: Notice that SE is the negative of the natural log of the H-K concentration unit with $\alpha = 1$.[6] so deciding between HHI and SE is equivalent to deciding on the perception parameter α and whether to satisfy *REP* or *SADD*. Further, since both HHI and SE are additively separable, they both satisfy *IND*, even though neither has an explicit consideration for the evenness of minority groups.

As discussed in section 1 and 2, existing units fall short as a measure of heterogeneity, due to either lacking in comparability between systems or not being descriptive of systems/not reflecting information from the minority groups. For a measure to be descriptive of the heterogeneity in systems using ψ_n and φ, it should satisfy the fundamental axioms as well as *IND*, *PPT*, and *CON*.

[4] 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:

$$C^n(x) = \phi^{-1}\left[\sum_{i=1}^{n} s_i \phi(s_i)\right].$$

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

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

[6] $-\ln\left(H^n_{\alpha=1}(S)\right) = -\sum_{g=1}^{G} P_i ln(P_i)$.

4 The Descriptive Units of Heterogeneity

Using the lessons learned from the other units of heterogeneity, I propose the *Descriptive Units of Heterogneeity*—a class of units that balances interpretability and comparability.

Let $n_1 \geq n_2 > 0$, $P_1 = \frac{n_1}{n_1+n_2+\cdots+n_G}$, and $\tilde{P}_g = \frac{n_g}{n_2+\cdots+n_G}$.

The Descriptive Units of Heterogeneity (DUH) of system S with $G \geq 2$ groups is defined as:

$$DUH(S) = \frac{ln(P_1)}{ln(G)} \cdot \left[\left(\sum_{g=2}^{G} \left| \tilde{P}_g - \frac{1}{G-1} \right|^p \right)^{\frac{1}{p}} - 1 \right].$$

Theorem: The *Descriptive Units of Heterogeneity* is the unique class of units, up to positive scalar multiplication, that satisfies *Scale Invariance*, *Group Symmetry*, *Independence*, *Principle of Diminishing Transfers*, *Principle of Proportional Transfers*, *Contractibility*, and uses ψ_p to incorporate the measure of evenness.

Notice that the parameter p controls how much evenness is reflected in DUH through ψ_p. As p increases, minority groups that are farther away from $\frac{1}{G-1}$ takes more weight. As $p \to \infty$,

$$d_p \to d_\infty = \sup_{g \in \{1,\ldots,G\}} \left\{ \left| P_g - \frac{1}{G-1} \right| \right\}.$$

Figure 2 shows how the progression of DUH changes with different p's. As p increases, the contribution of the evenness in minority takes less weight. When $p = 2000$, the effects of transfers between minority groups becomes negligible, making it look like P_1 dominates evenness in calculating DUH.

Figure 2: DUH with Different p

$p = 20$ $p = 200$ $p = 2000$

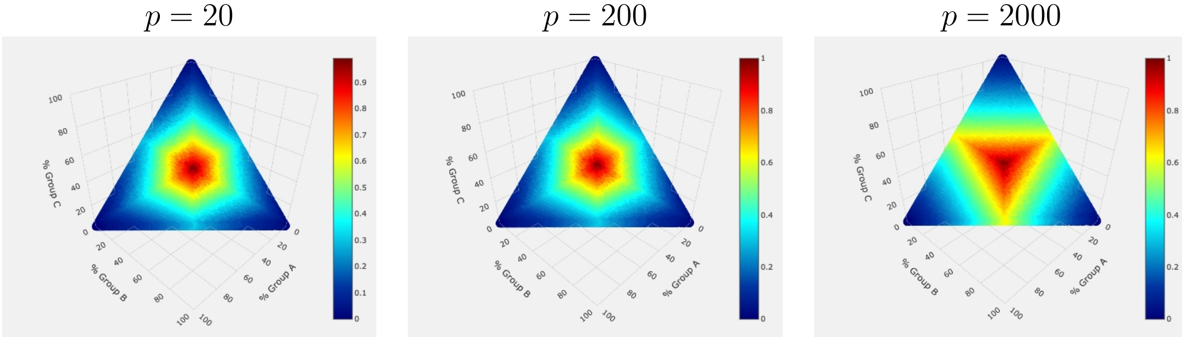


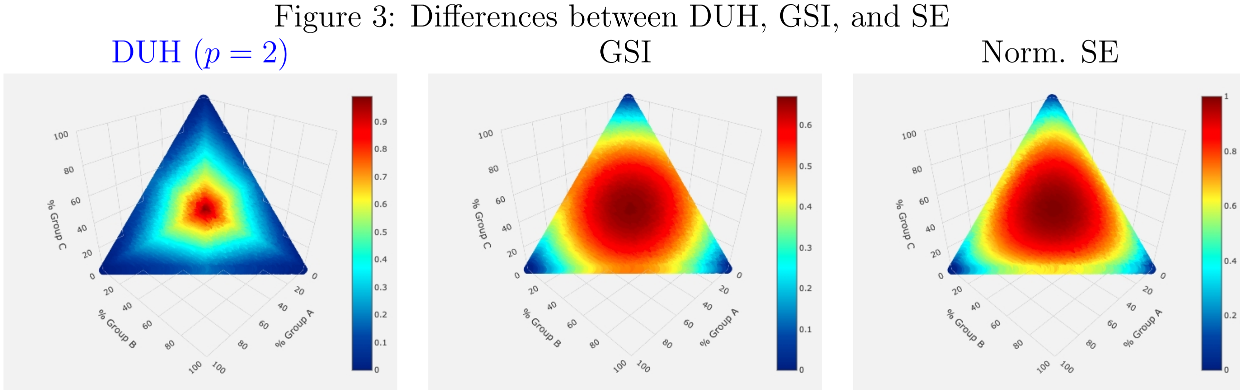
Table 1 outlines the axioms discussed and whether Gini, DUH, HHI, or SE satisfy them. Notice that DUH can be considered a refined Gini coefficient with generalized comparability across systems made up of discrete and unordered groups. DUH builds on the partial ordering of Gini coefficient to induce a total order that enables better comparisons across systems that do not satisfy the Lorenz criterion by further refining PDT adn incorporating ψ_p. DUH is characterized differently than concentration units because it focuses on the overall distribution without losing comparability.

Table 1: Measures and Axioms

Type	Axiom	Gini	DUH	HHI	SE
Fundamental	Type Symmetry	✓	✓	✓	✓
	Scale Invariance	✓	✓	✓	✓
	Principle of Diminishing Transfers	✓	✓	✓	✓
Characterization	Independence	×	✓	✓	✓
	Principle of Proportional Transfers	×	✓	×	×
	Expandability	×	×	✓	✓
	Contractibility	✓	✓	×	×
	Replication Principle	×	×	✓	×
	Shannon's Additivity	×	×	×	✓

Figure 3 compares DUH to the concentration units where $G = 3$. The tetrahedrons of each measure[7] below show how each measure changes as the distribution of groups become more heterogeneous. The centers of the triangles represent a system that is perfectly heterogeneous, and the vertices of the triangles represent systems that are perfectly homogeneous.

Figure 3: Differences between DUH, GSI, and SE



[7] SE is normalized to between 0 and 1 by dividing it by $ln(3)$.

5 Practical Uses of DUH

DUH, as a simple description of heterogeneity in systems, can be used in various contexts with discrete distributions over unordered groups in a system. This section provides several empirical examples where the strength of DUH is shown. Since HHI is a measure of concentration/homogeneity between 0 and 1, I used GSI (=1-HHI) here to make the comparisons simpler. Similarly, SE is normalized to be between 0 and 1 to make the comparisons simpler.

5.1 Reasonable Universe Θ

Recall that to measure the heterogeneity of any system, the system must first be thought of as a mapping from a universe of groups. Let $\Theta = \{\theta_1, \ldots, \theta_G\}$ be a **universe** of G distinct groups/categories. A system S is a mapping from Θ to $\mathbb{Z}_+^G$ such that $S = (n_1, \ldots, n_g \ldots, n_G)$ is a $1 \times G$ vector where n_g is a positive integer that represents the number of elements in θ_g in the system S with population n_S.

The implication of this paradigm is that the set Θ needs to be handled with care, because each $\theta \in \Theta$ must be similar/comparable to each other. Figures 4 and 5 illustrates this idea with a practical example. These two figures present racial composition of San Francisco MSA from 2007 to 2022 using ACS 1-year data (Ruggles et al. 2024). Figure 4 defines Θ as {White, Black, Other} while figure 5 splits up the *Other* group into 3 sub-groups, yielding $\Theta =$ {White, Black, Asian, Native America, Multi-Race}.

In figure 4, the heterogeneity of this system is somewhat stable due to the influence of the shrinkage in the white population and increase in the other population. The heterogeneity starts to decrease post 2019 when the white population became a minority group and the other population became a majority group. This change shows the importance of the axiom *SYM* which allows researchers to study heterogeneity as a distributional property free of labels. However, the story is different once Θ is redefined to further capture distributional changes in subgroups.

Figure 5 shows that when the Asian population and multi-race population is considered separately, heterogeneity actually increased post 2019, as the groups, at a glance, are proportionally growing. Such distributional changes is what *PPT* is designed to reflect.

When measuring heterogeneity in a system, one must realize the implications of choosing Θ. Determining the elements of Θ is a framing problem and is a judgement call by the researcher. Just as the use of Gini coefficient requires the *Lorenz Criterion*, the use of any units of heterogeneity requires justification of the reasonable groupings. In the examples here, the simple split of a subgroup changed the inference, serving as an excellent reason to why these units needs to be used with much care.

Keeping this in mind, let us consider examples of when DUH can be used and why it should be used. For simplicity, I will use the version of DUH where $p = 2$, and I would recommend others to do the same.

Figure 4: Comparisons between Different Units for Racial Heterogeneity

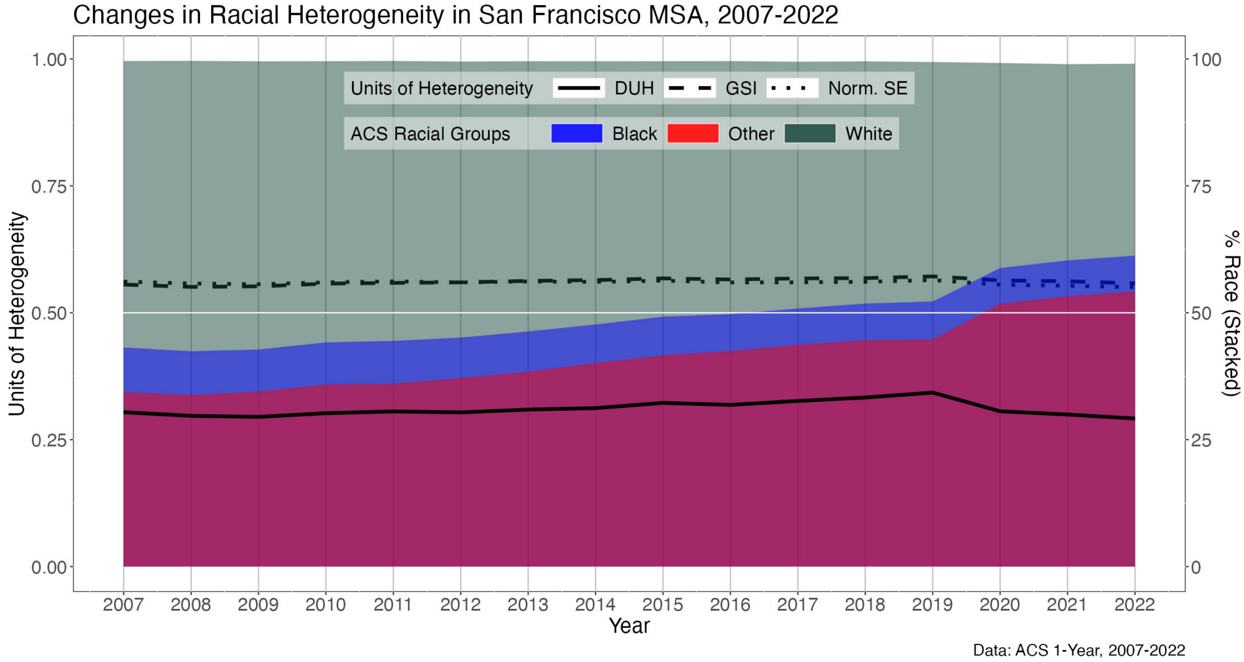
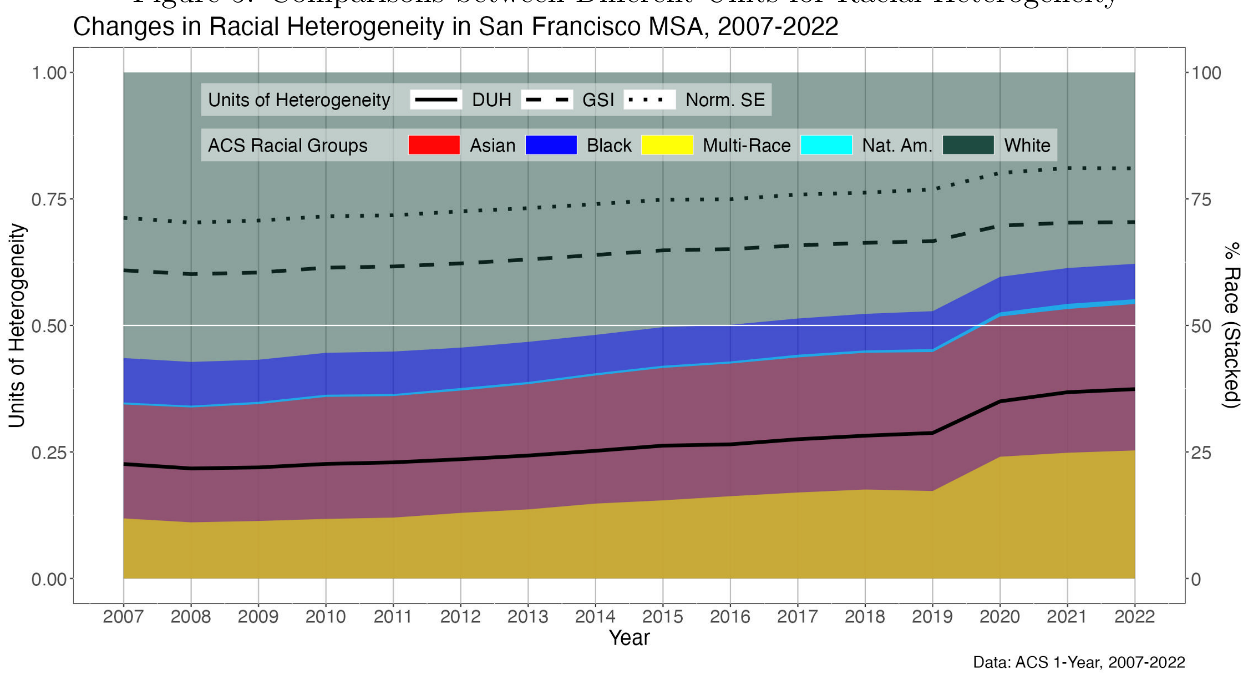


Figure 5: Comparisons between Different Units for Racial Heterogeneity



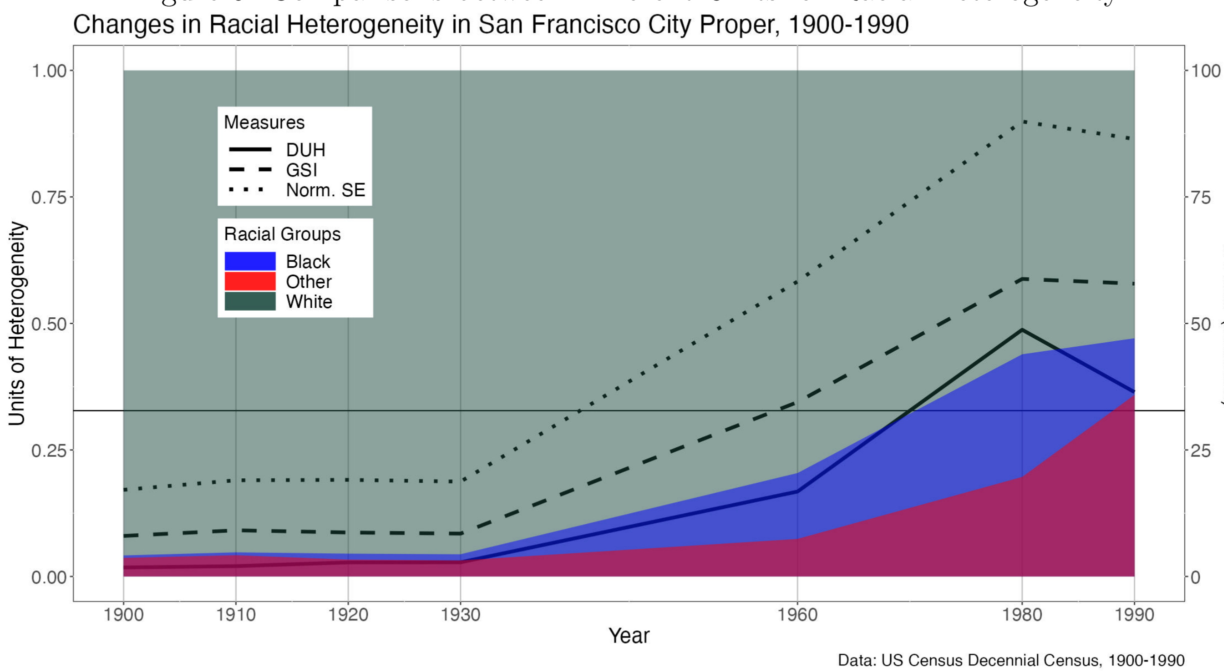
5.2 Examples

The examples here demonstrate how DUH can be useful for interpreting heterogeneity in different environments. The examples are arraigned to progress in the size of the reasonable universe of groups to show that DUH is sensitive only to the two components—size of majority and evenness of minority—and not to the number of groups.

5.2.1 Using DUH for Racial Heterogeneity

The first example uses DUH to measure racial heterogeneity when there are only 3 groups—White, Black, and Other—in the reasonable universe Θ. Figure 6 shows the progression of racial heterogeneity in San Francisco city proper from 1900-1990[8] using the decennial Census data from IPUMS USA (Ruggles et al. 2024).

Figure 6: Comparisons between Different Units for Racial Heterogeneity



From 1980 to 1990, the majority population (white) of San Francisco city proper decreased slightly, but the other population (mostly Asian) grew so much that it made the minority groups distributions much less even. In this case, GSI indicated only a slight the decrease in heterogeneity while the larger decrease in SE reflects more of this change in the minority group distribution. DUH on the other hand, follow generally the same trends as GSI and SE, yet it is able to reflect much more of the decrease in evenness in the minority distribution.

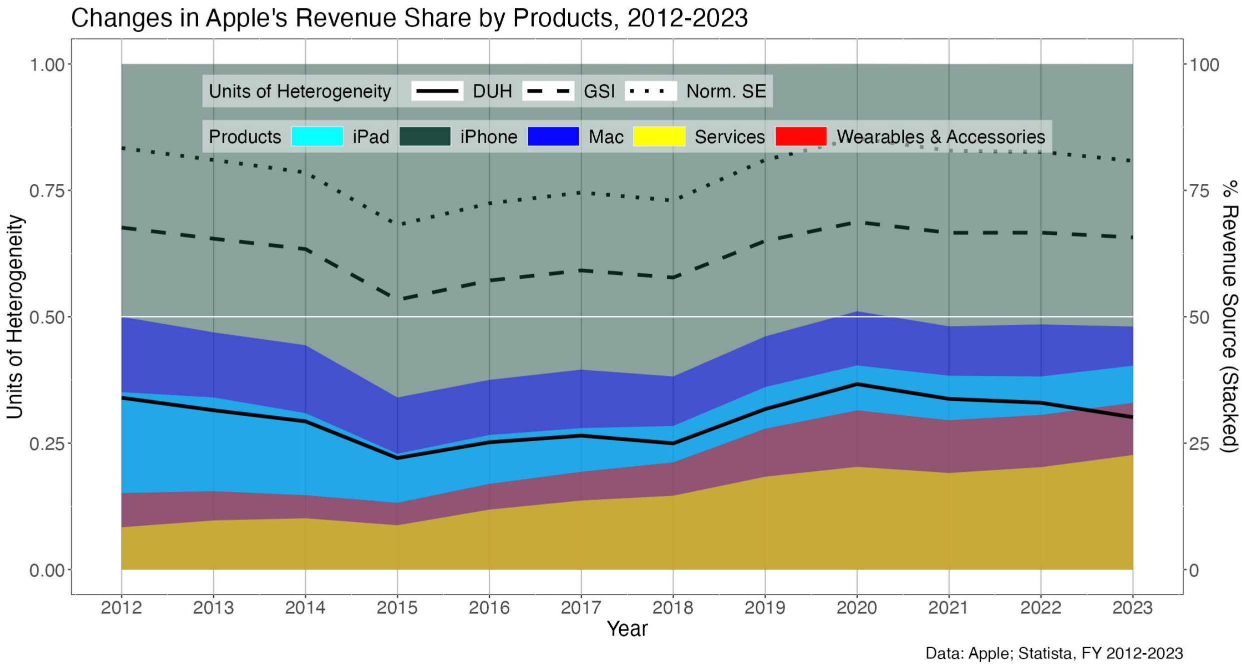
Recall that the main weakness in GSI and SE is that the size of the influence from changes in a group positively correlates with the size of the group. This example shows that DUH is able to dial back the correlation and reflect changes in the evenness of the minority groups.

[8]Due to Census coding of the inner-city variable, data is missing for 1940, 1950, and 1970.

5.2.2 Using DUH for Product Heterogeneity

The second example uses DUH to proxy how lightly a firm relies on specific products for its revenue. In this example, there are 5 groups—iPhone, iPad, Mac, Wearables & Accessories, and Services—in the reasonable universe for Apple's revenue, Figure 7 illustrates how DUH compares with GSI and DUH in a space that often utilize units of concentration using data on Apple's revenue source by products (Apple and Statista 2024). For the most part, the three units move in the same way. However, notice that from 2020 to 2023, Apple's revenue share for services as well as wearables (like apple watch) & accessories (like airpods) grew without diminishing the revenue share of iPhones. This decrease in the evenness in minority groups captured by a continuous and sizable decrease in DUH, while decreases in GSI in this period is limited.

Figure 7: Comparisons between Different Units for Product Heterogeneity



6 Summary and Discussion

Building on the axiomatizations of Gini, HHI, and SE, I uniquely characterize the set of units/indices/measures called the *Descriptive Units of Heterogeneity.* This set of units are simple to use, similar to existing units, and can fairly reflect changes in the evenness of minority groups. DUH provides a specific meaning to the sentence "S is x times more heterogeneous than S'" via the *Principle of Proportional Transfers.* Lastly, DUH can likely be extended as a measure for sorting by taking the mean squared differences between the DUH of a system and the DUH of the partitions of said system.

DUH is not meant to be a substitute of either HHI or SE. The improvement in reflecting the addition of a $0-$group and the changes in minority groups may not be a desirable property in cases where truly only information from large groups should be considered. In the end, I only hope that my pursuit of this new measure is guided by a reasonable and common motivation, and that future researchers, empirical or otherwise, can utilize this measure, or others similar to it, to find more insights in the evolution of heterogeneity in systems.

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.

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

Chakrabarti, C., Chakrabarty, I., et al. (2005). Shannon entropy: axiomatic characterization and application. *International Journal of Mathematics and Mathematical Sciences*, 2005:2847–2854.

Chakravarty, S. R. and Eichhorn, W. (1991). An axiomatic characterization of a generalized index of concentration. *Journal of Productivity Analysis*, 2:103–112.

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

James, D. R. and Taeuber, K. E. (1985). Measures of segregation. *Sociological Methodology*, 15:1–32.

Kolm, S.-C. (1976). Unequal inequalities. ii. *Journal of Economic Theory*, 13(1):82–111.

Kvålseth, T. O. (2022). Measurement of market (industry) concentration based on value validity. *Plos one*, 17(7):e0264613.

Marshall, A. W., Olkin, I., and Arnold, B. C. (1979). *Inequalities: theory of majorization and its applications*. Springer.

Nambiar, K., Varma, P. K., and Saroch, V. (1992). An axiomatic definition of shannon's entropy. *Applied Mathematics Letters*, 5(4):45–46.

Nunes, A., Trappenberg, T., and Alda, M. (2020). The definition and measurement of heterogeneity. *Translational Psychiatry*, 10(1):299.

Rothschild, M. and Stiglitz, J. E. (1969). Increasing risk: a definition and its economic consequences. *Cowles Foundation Discussion Paper*.

Rothschild, M. and Stiglitz, J. E. (1973). Some further results on the measurement of inequality. *Journal of Economic Theory*, 6(2):188–204.

Ruggles, S., Flood, S., Sobek, M., Backman, D., Chen, A., Cooper, G., Richards, S., Rodgers, R., and Schouweiler, M. (2024). IPUMS USA: Version 15.0. dataset.

Schwartz, J. and Winship, C. (1980). The welfare approach to measuring inequality. *Sociological Methodology*, 11:1–36.

Suyari, H. (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.

Appendix: Proofs

Definition 3: A function $\psi : \mathbb{R}_+^{G-1} \to \mathbb{R}_+$ is a measure of evenness in minority group distribution if it is of the following form:

$$p \in \mathbb{R}_+,\ \psi_p(S) = \psi\left(\tilde{P}_2, \dots, \tilde{P}_G\right) = 1 - \left(\sum_{g=2}^{G} \left|\tilde{P}_g - \frac{1}{G-1}\right|^p\right)^{\frac{1}{p}}.$$

Proposition 1: Consider an index $\Phi_p = \Phi(\varphi, \psi)$ that satisfies *SYM*, *INV*, and *IND*. Holding P_1 constant, if Φ is strictly increasing in ψ_p, then $\Phi(\varphi(P_1), \psi)$ satisfies the *Principle of Diminishing Transfers* if and only if $p > 1$.

Proof of Proposition 1: Consider two ordered systems $S = (P_1, \dots, P_g, P_{g+1}, \dots, P_G)$ and $S' = (P_1, \dots, P_g - c, P_{g+1} + c, \dots, P_G)$ where $c < \frac{P_g - P_{g+1}}{2}$. Define $\tilde{c} = \frac{c}{P_2 + \cdots + P_G}$. I want to show that $\psi(S) < \psi(S')$ and $\Phi(S) < \Phi(S')$, thus satisfying *PDT*.

Given S and S', we have

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

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

Observe that

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

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

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

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

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

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

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

□

Lemma 1: Any measure $\Phi(n_1, \dots, n_G)$ of system $S = (n_1, \dots, n_G)$ satisfies type symmetry if $(n_1, \dots, n_G)$ is a vector ordered such that $n_1 \geq n_2 \geq \dots \geq n_G$.

Proof of Lemma 1: The proof is trivial given that the groups are ordered by size and not the label of the groups. This is a convenient consequence of defining systems as mappings from the universe of groups to a vector of numbers.

Lemma 2: If an index $\Phi(\varphi, \psi)$ of heterogeneity satisfies *Scale Invariance* and *Principle of Proportional Transfers*, it is monotonically decreasing in P_1, and therefore a positive monotonic transformation of $\frac{1}{P_1}$.

Proof of Lemma 2:
Take any 2 systems of G groups $S = (n_1, n_2, \dots, n_G)$ and $S' = (n'_1, n'_2, \dots, n'_G)$ such that $\Phi(S) > \Phi(S')$ and that the $(n_2, \dots, n_G) = \tilde{S} = \lambda \cdot \tilde{S}' = \lambda \cdot (n'_2, \dots, n'_G)$, $\lambda \in \mathbb{R}_{++}$, then by *Scale Invariance*:

$$\Phi(n_1, n_2, \dots, n_G) > \Phi(n'_1, n'_2, \dots, n'_G) = \Phi\left(n'_1 \cdot \frac{n_S}{n'_S}, n'_2 \cdot \frac{n_S}{n'_S}, \dots, n'_G \frac{n_S}{n'_S}\right)$$

By the *Principle of Transfers*, since $n_1 + n_2 + \dots + n_G = n'_1 \frac{n_S}{n'_S} + n'_2 \frac{n_S}{n'_S} + \dots + n'_G \frac{n_S}{n'_S}$,

$$\Phi(S) > \Psi(S') \iff n_1 < n'_1 \cdot \frac{n_S}{n'_S} \iff P_1 < P'_1$$

□

Lemma 3: If an index $\Phi(\varphi, \psi)$ of heterogeneity satisfies *Scale Invaraince*, *Independence*, *Scale Invariance*, and *Principle of Proportional Transfers*, then φ and ψ must be multiplicatively separable.

Proof of Lemma 3:
Notice first that *Independence* trivially implies that φ and ψ must be separable. Take any system S with G groups. By the *Principle of Proportional Transfers*, it must be that

$\forall \alpha \in \left[1, \frac{n_1+\tilde{P}_2 \cdot n_1}{n-2+\tilde{P}_2 \cdot n_1}\right]$

$$\alpha \cdot \Phi(P_1, P_2, \ldots, P_G) = \Phi\left(P_1^\alpha, P_2 + \tilde{P}_2\left(P_1 - P_1^\alpha\right), \ldots, P_G + \tilde{P}_G\left(P_1 - P_1^\alpha\right)\right)$$
$$\iff \alpha\Phi(P_1, P_2, \ldots) = \Phi\left(P_1^\alpha, P_2', \ldots, P_G'\right) \iff \alpha = \frac{\varphi\left(P_1^\alpha\right)}{\varphi(P_1)}$$
where $\exists \lambda \in \mathbb{R}_{++} \, s.t. \, \lambda P_g = P_g', \forall g \in \{2, \ldots, G\}$

□

Proposition 2: If an index $\Phi(\varphi, \psi)$ of heterogeneity satisfies *Independence*, *Scale Invariance*, and *Principle of Proportional Transfers*, then it must be $\Phi = \varphi(P_1) \cdot \psi(\tilde{P}_2, \ldots, \tilde{P}_G)$ where $\varphi(P_1) = -c \cdot log_q(P_1)$, $c \in \mathbb{R}_{++}$

Proof of Proposition 2:
From the previous 2 lemmas, we know that $\varphi(P_1)$ must be a positive monotonic transformation of $\frac{1}{P_1}$ and that for α such that $P_1^\alpha > P_2 + \tilde{P}_2(P_1 - P_1^\alpha)$, we must have

$$\alpha = \frac{\varphi(P_1^\alpha)}{\varphi(P_1)}$$

Notice that the only positive monotonic transformation that would satisfy this is $log_q\left(\frac{1}{P_1}\right)$, up to a positive scalar multiplication. Further notice that any $log_q\left(\frac{1}{P_1}\right)$ can be rewritten as $\frac{ln\left(\frac{1}{P_1}\right)}{ln(q)}$, so it is equivalent to write $c \cdot ln\left(\frac{1}{P_1}\right)$. As such, $\varphi(P_1) = c \cdot ln\left(\frac{1}{P_1}\right)$, $c \in \mathbb{R}_{++}$ is the unique function, up to positive scalar multiplication, of majority proportions that can lead to $\Phi(\varphi, \psi)$ satisfying, *Independence*, *Scale Invariance*, and *Principle of Proportional Transfers*. □

Theorem 1: The *Descriptive Units of Heterogeneity* Φ defined as:

$$\Phi_p(n_1, \ldots, n_G) = -\frac{ln\left(P_1\right)}{ln(G)}\left[1 - \left(\sum_{g=2}^{G}\left(\tilde{P}_g - \frac{1}{G-1}\right)^p\right)^{\frac{1}{p}}\right]$$
$$= \frac{ln\left(P_1\right)}{ln(G)}\left[\left(\sum_{g=2}^{G}\left(\tilde{P}_g - \frac{1}{G-1}\right)^p\right)^{\frac{1}{p}} - 1\right]$$
$$\text{where } P_1 = \frac{n_1}{n_1 + \cdots + n_G}, \; \tilde{P}_g = \frac{n_g}{n_2 + \cdots + n_g}, \; p \in (1, \infty)$$

is the unique class of units that satisfy *Scale Invariance*, *Group Symmetry*, *Independence*, *Principle of Diminishing Transfers*, *Principle of Proportional Transfers*, *Contractibility*, and uses ψ_p to account for evenness in minority.

Proof of Theorem 1:
Propositions 1 and 2 combined implies that DUH satisfies *SYM*, *INV*, *PPT*, *IND*, and *CON*, but not necessarily *PDT*. I have only shown that DUH satisfies *PDT* when either φ or ψ_p is held constant, meaning I need to be careful with the functional form of $\varphi(P_1)$. To satisfy *PDT* overall, it must be that when there is a transfer from the majority group to the minority group, the increase in Φ through $\varphi(P_1)$ dominates the decrease in Φ through ψ_p in the case where evenness decreases as a result of the transfer.

Notice that to show Φ satisfies *PDT*, we only need to look at the extreme case where P_1 is close to 1 and $\psi = 1$. In this case, a simple transfer from n_1 to n_2 will decrease ψ_p the most. For simplicity, we will consider the case when $p = 2$ so that Φ is simply:

$$\Phi(n_1, \dots, n_G) = \frac{ln\,(P_1)}{ln(G)} \left[\sqrt{\sum_{g=2}^{G} \left(\tilde{P}_g - \frac{1}{G-1} \right)^2} - 1 \right]$$

Denote $n_2 + \cdots + n_G$ as $\tilde{n}_S$, a transfer of x from n_1 to n_2 when $\psi = 1$ can be written as:

$$\Phi_2 = ln \left(\frac{n_1 - x}{n_S} \right) \left[\sqrt{ \left(\frac{\frac{\tilde{n}_S}{G-1} + x}{\tilde{n}_S + x} \right) + (G-2) \left(\frac{\frac{\tilde{n}_S}{G-1}}{\tilde{n}_S + x} - \frac{1}{G-1} \right)^2 } - 1 \right]$$

Taking the derivative of this expression with respect to x, we have, $\forall x \in \left[0, \frac{(G-1)n_1 - \tilde{n}_S}{G}\right]$:

$$\frac{d}{dx}\Phi_2(n_1 - x, n_2 + x, n_3, \dots, n_G) = \frac{\sqrt{\frac{G-2}{G-1}} \left[\tilde{n}_S(x - n_1) ln \left(\frac{n_1 - x}{n_S} \right) + x(b + x) \right]}{(x - n_1)(x + \tilde{n}_S)^2} + \frac{1}{n_1 - x} > 0$$

$\square$